



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 122813

TO: David Lukton
Location: REM-3B75/3C70
Art Unit: 1653
Tuesday, May 25, 2004

Case Serial Number: 09/977349

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

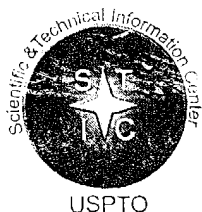
toby.port@uspto.gov

Search Notes

Dear Examiner Lukton,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
571-272-2507 Remsen E01 D86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:13:28 ; Search time 42.8108 seconds
(without alignments)
105.599 Million cell updates/sec

Title: US-09-977-349-1

Perfect score: 92

Sequence: 1 RQIKIWFEONRMMKWK 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A Geneseq_23Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	16	2	AAW45974 Cysteine
2	92	100.0	16	2	AAW33407 Peptide 4
3	92	100.0	16	2	AAW33410 D-form pe
4	92	100.0	16	2	AAW82958 Oestrogen
5	92	100.0	16	2	AAW56397 Preferred
6	92	100.0	16	2	AAW56397 Preferred
7	92	100.0	16	2	AAW71270 Antennape
8	92	100.0	16	2	AAW71316 Antennape
9	92	100.0	16	2	AAW30508 Drosophil
10	92	100.0	16	2	AAW91046 Internal
11	92	100.0	16	2	AAW52102 Peptide f
12	92	100.0	16	2	AAW00859 Peptide p
13	92	100.0	16	2	AAW13509 Signal se
14	92	100.0	16	3	AAW87920 Drosophil
15	92	100.0	16	3	AAW27060 Beta-cate
16	92	100.0	16	3	AAW93667 Peptide w
17	92	100.0	16	3	AAW67966 Carboxyl
18	92	100.0	16	3	AAW93551 Amino aci
19	92	100.0	16	3	AAW55818 Signal se
20	92	100.0	16	3	AAW71008 Drosophil
21	92	100.0	16	3	AAW51212 Antennape
22	92	100.0	16	3	AAW51167 Drosophil
23	92	100.0	16	3	AAW10343 Peptide A
24	92	100.0	16	3	AAW19251 Fragment
25	92	100.0	16	3	AAW93178 Protegrin
					AAW35694 Peptide a

26	92	100.0	16	3	AAW22025 Membrane
27	92	100.0	16	3	AAW29423 ANTPT pept
28	92	100.0	16	3	AAW03927 Internal
29	92	100.0	16	3	AAW93954 Peptide u
30	92	100.0	16	3	AAW29574 Antennape
31	92	100.0	16	3	AAW14785 Carrier m
32	92	100.0	16	3	AAW14785 Carrier m
33	92	100.0	16	3	AAW73091 Rheumatol
34	92	100.0	16	4	AAW60004 Internal
35	92	100.0	16	4	AAW70753 Cell memb
36	92	100.0	16	4	AAW02974 Protein t
37	92	100.0	16	4	AAW60671 Antennape
38	92	100.0	16	4	AAW06064 Drosophil
39	92	100.0	16	4	AAW49914 HIF-lalpa
40	92	100.0	16	4	AAW66996 Antennape
41	92	100.0	16	4	AAW00813 Fruit fly
42	92	100.0	16	4	AAW12205 Membrane
43	92	100.0	16	4	AAW78030 Peptide d
44	92	100.0	16	5	AAW78985 Cell pene
45	92	100.0	16	5	AAW78214 Amino aci
46	92	100.0	16	5	AAW15616 Drosophil
47	92	100.0	16	5	AAW15781 Antennape
48	92	100.0	16	5	AAW78345 Antennape
49	92	100.0	16	5	AAW68406 Translaca
50	92	100.0	16	5	AAW83153 Transduct
51	92	100.0	16	5	AAW53809 Penetrati
52	92	100.0	16	5	AAW25451 Drosophil
53	92	100.0	16	5	AAW75500 Fruit fly
54	92	100.0	16	5	AAW60447 Selective
55	92	100.0	16	5	AAW30763 E2F-1/cyc
56	92	100.0	16	5	AAW78912 Antennape
57	92	100.0	16	5	AAW81174 Antennape
58	92	100.0	16	5	AAW77234 Drosophil
59	92	100.0	16	5	AAW95801 Cell pene
60	92	100.0	16	5	AAW54754 Antennape
61	92	100.0	16	5	AAW47571 Penetrati
62	92	100.0	16	5	AAW48196 Drosophil
63	92	100.0	16	5	AAW23684 Fluoresce
64	92	100.0	16	6	AAW72298 ANTENNAP
65	92	100.0	16	6	AAW32768 DGI-3 spe
66	92	100.0	16	6	AAW35385 Antennape
67	92	100.0	16	6	AAW73390 Penetrati
68	92	100.0	16	6	AAW82532 Drosophil
69	92	100.0	16	6	AAW99502 Sequence
70	92	100.0	16	6	AAW19913 Antennape
71	92	100.0	16	6	AAW76121 Fruitfly
72	92	100.0	16	6	AAW56859 Antennape
73	92	100.0	16	6	AAW09581 Cell perm
74	92	100.0	16	6	AAW96963 Anti-infl
75	92	100.0	16	6	AAW84442 Penetrati
76	92	100.0	16	6	AAW07549 Fruitfly
77	92	100.0	16	6	AAW33891 Drosophil
78	92	100.0	16	6	AAW33889 Drosophil
79	92	100.0	16	6	AAW16665 Drosophil
80	92	100.0	16	6	AAW43125 Human mem
81	92	100.0	16	6	AAW70230 Membrane
82	92	100.0	16	6	AAW70225 Membrane
83	92	100.0	16	6	AAW79772 Penetrati
84	92	100.0	16	6	AAW09987 Antennape
85	92	100.0	16	6	AAW57028 Antennape
86	92	100.0	16	6	AAW82912 Antennape
87	92	100.0	16	6	AAW39279 Drosophil
88	92	100.0	16	6	AAW38246 Drosophil
89	92	100.0	16	7	AAW98598 Translaca
90	92	100.0	16	7	AAW39031 Human tra
91	92	100.0	16	7	AAW38685 Antennape
92	92	100.0	16	7	AAW38689 D-penetra
93	92	100.0	16	7	AAW42900 D-penetra
94	92	100.0	16	7	AAW42896 Antennape
95	92	100.0	16	7	AAW28156 Penetrati
96	92	100.0	16	7	AAW36650 Drosophil
97	92	100.0	16	7	AAW86319 Antennape
98	92	100.0	17	3	AAW83520 Peptide f

99 92 100.0 17 3 AAB13424 Synthetic
100 92 100.0 17 3 ADE14787 Carrier m

ALIGNMENTS

RESULT 1
AAW45974
ID AAW45974 standard; peptide; 16 AA.
XX AAW45974;
AC AAW45974;
XX
DT 01-JUL-1998 (first entry)
XX
DE Cysteine protease inhibiting peptide for preventing cell death.
XX
XX Neuronal cell death; neurodegenerative disorder; inhibition;
KW cysteine protease; cardiovascular; liver disease.
XX
OS Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1
FT - /note= "N-3-nitro-2-pyridyl-sulphenyl-Arg"

XX WO9735876-A1.
XX 02-OCT-1997.
XX
XX 04-MAR-1997; 97WO-US004159.
XX
XX 04-MAR-1996; 96US-00610220.
XX
XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Troy CM;
XX WPI; 1997-489561/45.
XX
PT New cysteine protease inhibiting peptide(s) for preventing cell death -
PT in cases of neurodegenerative, cardiovascular and liver diseases, and
PT their peptidomimetics, and general method for identifying enzyme
PT inhibiting peptides.
XX
PS Claim 8; Page 68; 112pp; English.

CC This sequence represents a specifically claimed peptide of the formula: V
CC -(AA1)n-Cys(V')-(AA2)m-V' (I), in which n and m = 0-5, totalling 2-5; if
CC n = 1, AA1 = Ala; if n = 2, (AA1)n = Gln-Ala; and if n = 3 or more, (AA1)n
CC = (X)p-Gln-Ala; X = any amino acid; p = 1-3, depending on value of n; if
CC m = 1, AA2 = Arg; if m = 2, (AA2)n = Arg-Gly; if m = 3 or more, (AA2)n =
CC Arg-Gly-(X)q; q = 1-3, depending on value of m; V, V' and V'', any or all
CC of which may be absent, = agent able to direct the compound to a specific
CC cell. The peptides are inhibitors of cysteine proteases, specifically
CC interleukin-1 beta converting enzyme (ICE). They inhibit death of cells,
CC particularly in humans, and can be used to treat neurodegenerative
CC diseases (e.g. ageing, Alzheimer's, Machado-Joseph, Parkinson's or
CC Huntington's diseases, multiple sclerosis, muscular dystrophy, stroke).
CC cardiovascular disease and liver disorders. The peptides should be more
CC specific than pseudosubstrate inhibitors

SQ Sequence 16 AA;
Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07; Mismatches 0; Gaps 0;
Matches 16; Conservative 0; Indels 0; Gaps 0;

OY 1 ROIKIWFQNRMRMKWK 16
| | | | | | | | | | | | | | | |
Db 1 ROIKIWFQNRMRMKWK 16

RESULT 2
AAW33407
ID AAW33407 standard; peptide; 16 AA.
XX
AC AAW33407;
XX

XX 27-AUG-2003 (revised)
DT 17-MAR-1998 (first entry)
XX
XX Peptide 43-58 of homeodomain Antp.

XX homeodomain; transcription factor; Antennapedia; Antp; vector;
KW transfection; hydrophobic.

XX Unidentified.

XX WO9712912-A1.

XX 10-APR-1997.

XX 04-OCT-1996; 96WO-FR001553.

XX 05-OCT-1995; 95FR-00011714.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Chassaing G, Prochiantz A;

XX WPI; 1997-226166/20.

XX New peptide(s) of high hydrophobic amino acid content - useful as vectors
PT for delivering peptides and nucleic acids to cells.

XX Claim 1; Page 7; 35pp; French.

XX New peptides are provided which are 16 amino acids long and which are
CC analogues of the peptide corresponding to residues 43-58 of the
CC Antennapedia transcription factor homeodomain (AntpHD). The peptides
CC contain 6-10 hydrophobic amino acids. They have the general formula: X1-
CC X2-X3-X4-X5-Trp-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16 or X16-X15-X14-X13-
CC X12-X11-X10-X9-X8-X7-Trp-X5-X4-X3-X2-X1 in which X1-X5 and X7-X16 are any
CC alpha-amino acids, provided that: (1) the peptide contains 6-10
CC hydrophobic amino acids; (2) X3 and X5 are not both Val; and (3) the
CC natural Antp 43-58 sequence ROIKIWFQNRMRMKWK (i.e. the present sequence)
CC is excluded. Specific examples of these peptides are given in AAW33408 -
CC AAW33416. The peptides are used as vectors for introducing into live
CC cells compounds which affect cell function, esp. peptides and nucleic
CC acids. They can cross cellular membranes and reach various cell
CC compartments. They are as effective as helix 3 of a homeodomain peptide.
CC (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 16 AA;

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07; Mismatches 0; Gaps 0;
Matches 16; Conservative 0; Indels 0; Gaps 0;

OY 1 ROIKIWFQNRMRMKWK 16
| | | | | | | | | | | | | | | |
Db 1 ROIKIWFQNRMRMKWK 16

RESULT 3
AAW33410
ID AAW33410 standard; peptide; 16 AA.
XX
AC AAW33410;
XX

DT 17-MAR-1998 (first entry)

DE D-form peptide 43-58 of homeodomain Antp.

XX homeodomain; transcription factor; Antennapedia; Antp; vector;

PI Nadler SG, Cleaveland JS, Blake J, Haffar OK;
 XX WPI; 1998-217028/19.
 XX
 PT Nuclear translocation inhibitor polypeptides - comprising signal sequence
 PT for delivery through the cytoplasmic membrane and at least 2 nuclear
 PT localisation sequences.
 XX
 XX Claim 5; Page 43; 69pp; English.
 XX
 CC Peptides AAW56397-99 represent preferred signal sequences of the
 CC invention. They are used to construct the nuclear translocation inhibitor
 CC polypeptides of the invention. Nuclear translocation inhibitor
 CC polypeptides comprise a signal sequence peptide capable of delivering the
 CC polypeptide through the cytoplasmic membrane into a cell, and at least 2
 CC nuclear localisation signals (NLSs). The polypeptides can be used to
 CC inhibit nuclear translocation of a cellular protein. In addition, since
 CC the nuclear translocation of certain cellular peptides is required for
 CC the host organism to mount an immune response, the polypeptide inhibitors
 CC are useful as immunosuppression agents. The polypeptides can therefore be
 CC used for the treatment of immune disorders including autoimmune diseases.
 CC The polypeptides can also be used for treating physical symptoms
 CC manifested by responses to allergens which can initiate a state of
 CC hypersensitivity, for the treatment of sepsis and in the prevention of
 CC septic shock, antiviral agents, tumour growth suppressors, and for
 CC transcriptionally modulating the expression of cellular genes
 XX
 SQ Sequence 16 AA;
 Query Match 100.0%; Score 92; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWFQNRRMKWKK 16
 Db 1 RQIKIWFQNRRMKWKK 16
 RESULT 6
 AAW71270
 ID AAW71270 standard; protein; 16 AA.
 AC AAW71270;
 XX
 DT 23-NOV-1998 (first entry)
 XX
 DE Antennapedia peptide for directing antisense oligonucleotides to a cell.
 KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;
 KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;
 KW Alzheimer's disease; dentatorubral; pallidolysial atrophy;
 KW Huntington's disease; Machado-Joseph disease; multiple sclerosis;
 KW muscular dystrophy; Parkinson's disease; senility;
 KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;
 KW trauma; antennapedia.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal protected with 3-nitro-2-pyridyl
 FT sulphenyl group (Npys)"
 FT
 XX W09838861-A1.
 XX
 XX 11-SEP-1998.
 XX
 XX 03-MAR-1998; 98WO-US004128.
 XX
 XX 03-MAR-1997; 97US-00810540.
 XX
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 XX Troy CM, Shelanski ML;
 XX WPI; 1998-506333/43.
 XX

PI Troy CM, Shelanski ML;
 XX WPI; 1998-506333/43.
 XX
 PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to
 PT inhibit neurodegenerative disorder associated with e.g. ageing,
 PT Alzheimer's, Huntington's or Parkinson's disease.
 XX
 XX Claim 5; Page 39; 60pp; English.
 XX
 CC AAW71270 and AAW71315-16 represent antennapedia peptides which are used
 CC to, direct the antisense oligonucleotides (AAV54973-74) of the invention
 CC to a cell. The antisense oligonucleotides are used to inhibit cell death
 CC mediated by withdrawal of a trophic factor. AAV54973 inhibits the
 CC expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of
 CC a human Ich-1L gene. The oligonucleotides are used to inhibit neuronal
 CC cell death, especially for treatment of neuronal cell death caused by
 CC e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease,
 CC dentatorubral and pallidolysial atrophy, Huntington's disease, Machoado-
 CC Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's
 CC disease, senility, spinocerebellar ataxia type I, spinobulbar muscular
 CC atrophy, stroke or trauma
 XX
 SQ Sequence 16 AA;
 Query Match 100.0%; Score 92; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWFQNRRMKWKK 16
 Db 1 RQIKIWFQNRRMKWKK 16
 RESULT 7
 AAW71316
 ID AAW71316 standard; protein; 16 AA.
 XX
 AC AAW71316;
 XX
 DT 23-NOV-1998 (first entry)
 XX
 DE Antennapedia peptide for directing antisense oligonucleotides to a cell.
 KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;
 KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;
 KW Alzheimer's disease; dentatorubral; pallidolysial atrophy;
 KW Huntington's disease; Machado-Joseph disease; multiple sclerosis;
 KW muscular dystrophy; Parkinson's disease; senility;
 KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;
 KW trauma; antennapedia.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal protected with 3-nitro-2-pyridyl
 FT sulphenyl group (Npys)"
 FT
 XX W09838861-A1.
 XX
 XX 11-SEP-1998.
 XX
 XX 03-MAR-1998; 98WO-US004128.
 XX
 XX 03-MAR-1997; 97US-00810540.
 XX
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX
 XX Troy CM, Shelanski ML;
 XX WPI; 1998-506333/43.
 XX

PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to
PT inhibit neurodegenerative disorder associated with e.g. ageing,
PT Alzheimer's, Huntington's or Parkinson's disease.

XX Disclosure; Page 14; 60pp; English.

XX AAW71270 and AAW71315-16 represent antennapedia peptides which are used
CC to, direct the antisense oligonucleotides (AAV54973-74) of the invention
CC to a cell. The antisense oligonucleotides are used to inhibit cell death
CC mediated by withdrawal of a trophic factor. AAV54973 inhibits the
CC expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of
CC a human Ich-IL gene. The oligonucleotides are used to inhibit neuronal
CC cell death, especially for treatment of neuronal cell death caused by
CC e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease, Machado-
CC Jostorubral and pallidolysual atrophy, Huntington's disease, Parkinson's
CC Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's
CC disease, senility, spinocerebellar ataxia type I, spinobulbar muscular
CC atrophy, stroke or trauma

XX Sequence 16 AA;

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 8

AAW30508
ID AAW30508 standard; peptide; 16 AA.

XX AC AAW30508;

XX DT 26-OCT-1998 (first entry)

XX DE Drosophila membrane translocation sequence.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
XX cell proliferation; cardiovascular cell; re-tenosis; tumour;
XX surgical stent; therapy; membrane translocation; antennapedia protein.

XX OS Drosophila melanogaster.

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB003506.

XX PR 20-DEC-1996; 96GB-00026589.

XX (PROL-) PROLIFIX LTD.

XX La Thangue NB, Bandara LR;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells, e.g.
PT for preventing re-stenosis.

XX Disclosure; Page 5; 55pp; English.

XX This polypeptide comprises a membrane translocation sequence derived from
CC the Drosophila melanogaster antennapedia protein. Such membrane
CC translocation sequences are useful in directing entry of a polypeptide
CC into a cell. Polypeptides of the invention (see AAW30504-07) are derived
CC from the DEF box region (see AAW30501) of transcription factor Dp1. They
CC act as antagonists of the heterodimerisation of a Dp protein with an E2F
CC protein, and can be used to induced apoptosis, specifically in tumour and

CC cardiovascular cells, e.g. for preventing restenosis. A claimed fusion
CC protein comprises a DEF box peptide and the Drosophila antennapedia
CC protein membrane translocation sequence

XX Sequence 16 AA;

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 9

AAW91046

ID AAW91046 standard; peptide; 16 AA.

XX AC AAW91046;

XX DT 24-MAR-1999 (first entry)

XX DE Internalization sequence associated with cadherin modulating agents.

XX Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
XX cadherin-mediated function; demyelinating neurological disease;
XX multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
XX central nervous system; apoptosis induction; cadherin-expression cell;
XX pregnancy prevention; vasopermeability; synaptic stability; diabetes;
XX rheumatoid arthritis; allergic response; learning; antennapedia protein;
XX antibody-mediated graft rejection; internalization sequence; memory.

XX OS Synthetic.

XX PN WO9845319-A2.

XX PD 15-OCT-1998.

XX PF 14-APR-1998; 98WO-CA000322.

XX PR 10-APR-1997; 97US-0043361P.

XX (UYWC-) UNIV MCGILL.

XX PI Blaschuk OW, Gour BJ;

XX WPI; 1999-024009/02.

XX New catenin modulating agents - comprising peptides having a sequence HAV
PT or analogues or antibodies, used for modulating cadherin-mediated
PT functions.

XX Claim 16; Page 78; 106pp; English.

XX The present sequence represents antennapedia protein derived
CC internalization sequence associated with cadherin modulating agents.
CC These agents are used for modulating cadherin-mediated functions. They
CC can be used for disrupting interaction between alpha-catenin and beta-
CC catenin in a cell, inhibiting cell adhesion, e.g. between epithelial
CC cells, endothelial cells, neural cells, tumour cells and lymphocytes, for
CC treating a demyelinating neurological disease, e.g. multiple sclerosis,
CC for reducing unwanted cellular adhesion in a mammal, for enhancing the
CC delivery of a drug through the skin of a mammal, for enhancing the
CC delivery of a drug to a tumour in a mammal, for treating cancer in a
CC mammal, for inhibiting angiogenesis in a mammal, for enhancing drug
CC delivery to the central nervous system of a mammal, for inducing
CC apoptosis in a cadherin-expression cell, for modulating the immune system
CC of a mammal, for preventing pregnancy in a mammal, for increasing
CC vasopermeability in a mammal, or for inhibiting synaptic stability in a
CC mammal. In particular they can be used for treating diabetes, rheumatoid
CC arthritis, allergic responses, antibody-mediated graft rejection or for
CC stimulating learning and memory

```

XX SQ Sequence 16 AA;
Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
DB 1 RQIKWIFQNRMKWK 16

RESULT 11
AAY00859
ID AAY00859 standard; peptide; 16 AA.
XX AC AAY00859;
XX DT 20-MAY-1999 (first entry)
XX DE Peptide pAntp(43-58) used in membrane-permeable construct.
XX KW Membrane-permeable construct; lipid membrane; membrane transport;
XX KW oligonucleotide delivery; cancer therapy; signal transduction; inhibitor;
XX KW gene therapy; transcription; translation; expression; replication.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 16
XX FT /note= "amidated"
XX PN WO9905302-A1.
XX PD 04-FEB-1999.
XX PF 16-JUL-1998; 98WO-US014761.
XX PR 24-JUL-1997; 97US-0053678P.
XX PX (PEKE ) PERKIN-ELMER CORP.
XX PI Langel U, Bartfai T, Pooga M, Valkna A, Saar K, Hallbrink M;
XX PF; 1999-142952/12.
XX New membrane-permeable constructs - comprise a peptide linked by a labile
XX bond to a nucleic acid analogue capable of hybridising with an
XX intracellular polynucleotide.
XX PS Disclosure; Page 26; 60pp; English.
XX CC This sequence represents a peptide used in the construct of the
XX invention. The construct is a membrane-permeable construct for transport
XX across a lipid membrane, which comprises: (a) a nucleic acid analogue
XX capable of hybridising with an intracellular polynucleotide (PN); (b) a
XX peptide; and (c) a labile bond linking the nucleic acid analogue and the
XX peptide. The membrane-permeable constructs can be used for delivery of
XX oligonucleotides, nucleic acids and nucleic acid analogues into cells.
XX They can be used for e.g. cancer therapy, signal transduction studies,
XX identifying new intracellular drug targets or gene therapy. They can also
XX be used for selectively inhibiting DNA transcription, RNA translation,
XX RNA or DNA expression, DNA replication, or an DNA or RNA regulatory
XX function of preselected DNA or RNA sequences in a living cell
XX SQ Sequence 16 AA;
Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
DB 1 RQIKWIFQNRMKWK 16

RESULT 12
AAY13509
ID AAY13509 standard; peptide; 16 AA.

```

XX AAY13509;
 XX 30-JUL-1999 (first entry)
 XX Signal sequence of antennapedia.
 XX Fusion protein; calpastatin; calpain; platelet aggregation; arthritis;
 KW hypoxia; erythrocyte sickling; sickle cell; HIV provirus; NF-kappaB;
 KW inflammation; asthma; immune response; restenosis; myocardial infarction;
 KW cancer; calpastat; antennapedia.
 XX Mammalia.
 XX WO9922756-A1.
 XX 14-MAY-1999.
 XX 04-NOV-1998; 98WO-US023526.
 XX 04-NOV-1997; 97US-00964302.
 XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
 XX Potter DA, Skolnik PR;
 XX WPI; 1999-326923/27.
 XX Fusion protein of signal sequence and calpastatin.
 XX Disclosure; Page 5; 46pp; English.
 XX The invention features fusion proteins that contain a calpastatin peptide
 CC and a signal sequence capable of delivering the fusion protein into a
 CC eukaryotic cell. The fusion protein is used for the inhibition of calpain
 CC in a cell. The fusion proteins are specifically used: (a) to prevent
 CC aggregation and degradation of platelets (e.g. during storage); (b) to
 CC inhibit hypoxia-induced sickling of erythrocytes (during storage,
 CC facilitating subsequent transfusion of autologous cells for treatment of
 CC sickle cell crises); and (c) to inhibit activation of human immune
 CC deficiency virus provirus in infected cells (or similarly for other
 CC viruses regulated by NF-kappaB). Other disclosed uses are: to treat or
 CC prevent inflammation (e.g. arthritis or asthma), unwanted immune
 CC responses (e.g. transplant rejection), restenosis (associated with
 CC angioplasty), cancer, subarachnoid hemorrhage, vasospasm, muscular
 CC dystrophy, cataracts and traumatic birth injury; to prevent spread of
 CC platelets on surfaces (e.g. when applied to the surface of stents,
 CC catheters etc.); to reduce coronary thrombosis in by-pass surgery and
 CC angioplasty; to treat myocardial infarction, or to prevent progression of
 CC infarction (myocardial or cerebral). The fusion protein has a reversible
 CC inhibitory effect and enters cells easily. It allows platelets to be
 CC stored cold with reduced change in shape
 XX Sequence 16 AA;
 XX Query Match 100.0%; Score 92; DB 2; Length 16;
 XX Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 RQIKWFOQRNMKWK 16
 XX 1 RQIKWFOQRNMKWK 16
 XX
 XX RESULT 13
 XX AAY87920
 XX ID AAY87920 standard; peptide; 16 AA.
 XX AC AAY87920;
 XX 11-SEP-2000 (first entry)
 XX Drosophila sp. antennapedia protein fragment.

XX Rad51; cell proliferation; rad51TR1-131; rad51K-A134; immunosuppressive;
 KW cytosstatic; antiinflammatory; antiproliferative; programmed cell death;
 KW treatment; autoimmune disorder; inflammation; cancer; graft rejection;
 KW proliferative disorder; hyperproliferative disorder; antennapedia.
 XX Drosophila sp.
 XX US6057104-A.
 XX 02-MAY-2000.
 XX 05-NOV-1997; 97US-00964614.
 XX 05-NOV-1996; 96US-00758280.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Hasty P;
 XX WPI; 2000-349568/30.
 XX Mutant forms of mammalian Rad51 proteins and polypeptides that disrupt
 PT cell proliferation and promote programmed cell death, encoded by specific
 PT nucleotides.
 XX Disclosure; Col 31-32; 22pp; English.
 XX This invention describes novel truncated and altered forms of Rad51
 CC products encoded by rad51TR1-131 and rad51K-A134. The products of the
 CC invention have immunosuppressive, cytostatic, antiinflammatory and
 CC antiproliferative activity. Altered and mutant forms of Rad51 are useful
 CC for inhibiting cell proliferation and for promoting programmed cell
 CC death. Therapeutic agents, factors or compounds capable of disrupting the
 CC essential processes mediated by or associated with normal Rad51 or Rad52
 CC are useful for treating diseases like autoimmune disorders, inflammation,
 CC cancer, graft rejection and other proliferative and hyperproliferative
 CC disorders. This sequence represents a fragment of the Drosophila
 CC antennapedia protein which is described in the method of the invention
 XX Sequence 16 AA;
 XX Query Match 100.0%; Score 92; DB 3; Length 16;
 XX Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 RQIKWFOQRNMKWK 16
 XX 1 RQIKWFOQRNMKWK 16
 XX
 XX RESULT 14
 XX AAB27060
 XX ID AAB27060 standard; peptide; 16 AA.
 XX AC AAB27060;
 XX 15-FEB-2001 (first entry)
 XX Beta-catenin derived internalisation moiety SEQ ID NO: 47.
 XX Beta-catenin; cadherin-mediated intercellular adhesion;
 KW cell differentiation; modulating agent; hair loss; skin exfoliation;
 KW internalisation moiety; flanking sequence; transcription; hearing loss.
 XX Drosophila melanogaster.
 XX WO200053632-A1.
 XX 14-SEP-2000.
 XX 07-MAR-2000; 2000WO-CA000222.

PR 09-MAR-1999; 99US-00265107.
 XX (UYMC-) UNIV MCGILL.
 PA Blaschuk OW, Gour BU;
 FI WPI; 2000-594308/56.
 XX Stimulating beta-catenin mediated gene expression, cellular
 PT differentiation and hair growth, involves contacting cells with
 PT modulating agent capable of inhibiting interaction between alpha and beta
 PT catenin.
 XX
 XX Claim 8; Page 38; 77pp; English.
 XX The present invention is concerned with methods of modulating the amount
 CC of free beta-catenin in the cell, and methods of stimulating the
 CC expression of genes involved in cellular differentiation, the
 CC transcription of which is under the control of beta-catenin. The peptides
 CC given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351 can be used
 CC as modulating agents which interrupt the interaction between alpha and
 CC beta catenin, causing increased levels of the latter and stimulating the
 CC activation of beta-catenin mediated transcription. This can be used to
 CC stimulate cell differentiation, which can then be used to promote hair
 CC growth and skin exfoliation. This latter is particularly useful in the
 CC improvement of photodamaged skin and to minimise wrinkles. The modulating
 CC peptide can also be used to reduce hearing loss resulting from inner ear
 CC disorders such as hyperacusis and tinnitus
 XX
 XX Sequence 16 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
 |||||
 Db 1 RQIKWQNRRMKWKK 16
 |||||

RESULT 15
 AAY93667
 ID AAY93667 standard; peptide; 16 AA.
 AC AAY93667;
 XX
 XX 25-SEP-2000 (first entry)
 DT
 XX Peptide which may be linked to anticancer agents.
 XX Anticancer agent; cancer cell; resistance; P-glycoprotein pump; cancer.
 XX Unidentified.
 OS
 XX WO200032237-A1.
 PN
 XX 08-JUN-2000.
 PD
 XX 26-NOV-1999; 99WO-FR002939.
 PF
 XX 30-NOV-1998; 98FR-00015073.
 XX (SYNT-) SYNT:EM SA.
 PA
 XX Tansamani J, Kaczorek M, Colin De Verdier A;
 PI WPI; 2000-412166/35.
 XX New composition useful for cancer treatment and prevention, contains
 PT anticancer agent and peptide vector that transports agent into cells.
 XX Disclosure; Page 8; 34pp; French.
 PS
 XX

CC The specification describes a pharmaceutical composition, which comprises
 CC at least one anticancer agent associated with at least one peptide that
 CC can transport it into cancer cells and which inhibits development of
 CC resistance to the anticancer agent. By using the peptide as a vector for
 CC delivery of the anticancer agent, mechanisms that cause cancer cells to
 CC become resistant to the agent, particularly the P-glycoprotein pump, are
 CC avoided. Also, peptides are easily produced by chemical synthesis, can be
 CC coupled easily to the agent, cross mammalian cell membranes rapidly by a
 CC passive mechanism (no receptors required), and are non-toxic and non-
 CC lytic. The compositions are used to treat cancer. The present sequence
 CC represents a peptide which may be linked to the anticancer agents of the
 CC invention
 XX
 XX Sequence 16 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
 |||||
 Db 1 RQIKWQNRRMKWKK 16
 |||||

RESULT 16
 AAY67966
 ID AAY67966 standard; peptide; 16 AA.
 XX
 AC AAY67966;
 XX
 XX 05-APR-2000 (first entry)
 DT
 XX Carboxyfluorescein cell permeant peptide #3.
 DE
 XX Kaposi syndrome; fibroblast growth factor; signal peptide; PNA;
 KW peptide nucleic acid; cell permeability; intracellular delivery;
 XW gene therapy; cancer.
 XX
 XX Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= 'labelled with carboxyfluorescein'
 FT
 FT
 XX WO964449-A2.
 XX
 XX 16-DEC-1999.
 PD
 XX 10-JUN-1999; 99WO-GB001848.
 PF
 XX 10-JUN-1998; 98GB-00012376.
 PR
 XX 10-JUL-1998; 98GB-00014888.
 XX
 XX (UYBE-) UNIV QUEENS BELFAST.
 XX
 XX Nelson J, Harriott P, Wallace A;
 PI WPI; 2000-097517/08.
 DR
 XX New cell permeable signal peptides, useful for intracellular delivery of
 PT a molecule.
 PT
 XX Example 1; Page 29; 33pp; English.
 PS
 XX The present invention describes a cell permeable peptide comprising at
 CC least the hydrophobic core of a signal peptide (or analogue) containing
 CC at least 1 additional positively charged amino acid (or analogue). The
 CC peptides are useful for the intracellular delivery of molecules,
 CC especially peptide nucleic acids to in vivo targets. The peptides are
 CC useful in commercial drug-delivery systems, in e.g. gene therapy, cancer
 CC therapy and anti-infectious agent therapy. The peptides facilitate
 CC biochemical and molecular biological research. The modified peptides
 CC facilitate intracellular delivery of any cell-impermeable substances and

CC improve delivery into low permeability cells. Delivery into sub-
 CC compartments can be achieved by modifying the signal peptides. The
 CC present sequence represents a peptide used in the exemplification of the
 CC present invention

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFOQRNMKWK 16
 |||||
 Db 1 RQIKWFOQRNMKWK 16

RESULT 17

AA93551
 ID AAY93551 standard; peptide; 16 AA.

XX AC AAY93551;

XX DT 25-SEP-2000 (first entry)

XX DE Amino acid sequence of a synthetic protein transduction domain.

XX KW Protein transduction system; protein transduction domain;
 XX KW cytotoxic domain; pathogen infection; retroviral infection;
 XX KW plasmoidal infection; cancer; prostate cancer.

XX OS Synthetic.

XX WO200034308-A2.

XX PD 15-JUN-2000.

XX PF 10-DEC-1999; 99WO-US029289.

XX PR 10-DEC-1998; 98US-0111701P.

XX PA (UNIV) UNIV WASHINGTON.

XX PI Dowdy SF;

XX WPI; 2000-431269/37.

XX PT Protein transduction system for treating cancer and pathogenic infections
 XX PT has a fusion protein comprising a protein transduction domain covalently
 XX PT linked to a cytotoxic domain.

XX PS Example 10; Page 71; 127pp; English.

XX CC AAY93542-51 represent synthetic protein transduction domains, which are
 CC used in the protein transduction system of the invention. The
 CC specification describes a protein transduction system, which comprises a
 CC fusion protein. This fusion protein has a covalently linked protein
 CC transduction domain and cytotoxic domain. The system is useful for
 CC treating pathogen infection in mammals, infections such as those caused
 CC by CMV, HSV-1, HCV, KSHV, yellow fever virus, flavivirus or rhinovirus,
 CC retroviral infections such as HIV-1, HIV-2, HTLV-3 and/or LAV, plasmoidal
 CC infections associated with P.faciiparum, P.vivax, P.ovale, P.malariae. It
 CC is also useful for treating cancer, especially prostate cancer

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFOQRNMKWK 16
 |||||
 Db 1 RQIKWFOQRNMKWK 16

RESULT 18

AA55818
 ID AAY55818 standard; peptide; 16 AA.

XX AC AAY55818;

XX DT 06-MAR-2000 (first entry)

XX DE Signal sequence for delivering polypeptide inhibitor.

XX KW Cellular protein; nuclear translocation; nuclear localization signal;
 XX KW immunosuppressant; immune response; viral infection; immune disorder;
 XX KW rheumatoid arthritis; multiple sclerosis; juvenile-onset diabetes; NLS;
 XX KW systemic lupus erythematosus; SLE; asthma; sepsis; tumor growth;
 XX KW fibroblast growth factor; signal sequence.

XX OS Synthetic.

XX WO9957138-A1.

XX PD 11-NOV-1999.

XX PF 26-APR-1999; 99WO-US008984.

XX PR 04-MAY-1998; 98US-00072429.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Nadler SG;

XX WPI; 2000-062141/05.

XX PT Novel composition comprising a nuclear localization signal and an
 XX PT inhibitor of nuclear translocation, useful for preventing transplanted
 XX PT organ or tissue rejection.

XX PS Disclosure; Page 14; 68pp; English.

XX CC The invention provides a composition comprising an isolated polypeptide
 CC (I) (polypeptide inhibitor of a cellular protein nuclear translocation)
 CC which comprises a signal sequence peptide capable of delivering (I)
 CC through a cytoplasmic membrane into the cell; and a nuclear localization
 CC signal sequence (NLS) present along with an immunosuppressant. (I) or its
 CC derivatives provide useful tools for introducing an exogenous polypeptide
 CC comprising an NLS into an intact cell to inhibit nuclear translocation of
 CC a cellular protein, for studying the role of nuclear translocation in the
 CC regulation of cellular processes. See AAY55812 for detailed uses of (I)
 CC and compositions containing (I). The present sequence represents a signal
 CC sequence for delivering the (I) through the cytoplasmic membrane into the
 CC cell

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFOQRNMKWK 16
 |||||
 Db 1 RQIKWFOQRNMKWK 16

RESULT 19

AA71008
 ID AAY71008 standard; peptide; 16 AA.

XX AC AAY71008;

XX DT 29-AUG-2000 (first entry)

XX DE Drosophila antennapedia (ANT) transport peptide.

KW Phospholamban; PLB; sarcoplasmic reticulum Ca 2+ ATPase; SERCA2a;
 XW cardiomyocyte; transport peptide; penetratin; cargo peptide; inhibitor;
 KW contractilin; cardiac contractility; cardiac disease; antennapedia; ANT;
 XX treatment; cardiant; heart failure; myocardial dysfunction; fruit fly.
 OS Drosophila sp.
 XX WO200025804-A2.
 PN 11-MAY-2000.
 PD 02-NOV-1999; 99WO-US025692.
 XX 02-NOV-1998; 98US-0106718P.
 PF 27-JUL-1999; 99US-0145883P.
 XX (REGC) UNIV CALIFORNIA.
 PA Chien K, Dillman W, Minamisawa S, He H, Hoshijima M, Meyer M;
 PI Scott C, Wang Y, Silverman GJ;
 XX WPI; 2000-365393/31.
 DR Treating cardiac diseases, e.g. heart failure or myocardial dysfunction
 XX comprises enhancing cardiac contractility by inhibiting interaction
 PT between phospholamban and sarcoplasmic reticulum calcium 2+ adenosine
 PT triphosphatase.
 XX Example 4; Page 50; 56pp; English.
 PS The patent discloses a method for the treatment of heart failure, using
 XX small peptide complexes and recombinant proteins, that induces
 CC phospholamban (PLB) deficiency and inhibits the interaction between PLB
 CC and sarcoplasmic reticulum Ca 2+ ATPase (SERCA2a) within cardiomyocytes.
 CC The peptide complex comprises of transport peptide like penetratin and
 CC cargo peptide selected from mutant PLB, native PLB or antibody against
 CC PLB protein (contractilin). Penetratin-PLB peptide functions as a
 CC dominant negative inhibitor of PLB-SERCA2a interaction, enhances cardiac
 CC contractility and reduces blood pressure. This method is useful for the
 CC treatment of cardiac disease e.g. heart failure and myocardial
 CC dysfunction. The present amino acid sequence is the transport peptide
 CC composed of residues 43-58 of antennapedia (ANT), a drosophila
 CC transcription factor protein. This peptide can be used to transport
 CC therapeutic agents across a cell membrane into the cytoplasm and nucleus
 CC of cardiomyocytes
 XX SQ Sequence 16 AA;
 Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 1 RQIKIWQNRRMKWKK 16
 |||||
 RESULT 20
 AAY51212
 ID AAY51212 standard; peptide; 16 AA.
 XX AC AAY51212;
 XX 04-APR-2000 (first entry)
 DT Antennapedia protein homeodomain peptide fragment 1.
 XX Antennapedia; homeodomain; fluorogenic; fluorescence; fluorophore;
 KW fluorescence resonance energy transfer; FRET; detection.
 XX Unidentified.
 OS EP969284-A1.
 XX PN

XX 05-JAN-2000.
 PD 01-JUL-1999; 99EP-00112544.
 PF 02-JUL-1998; 98DE-01029495.
 XX (PAYS/) PAYSAN J.
 PA (ANTZ/) ANTZ C.
 XX Paysan J, Antz C;
 PI WPI; 2000-099447/09.
 DR Fluorogenic conjugate for intracellular fluorescence labeling, especially
 XX for performing fluorescence resonance energy transfer assays in living
 PT cells.
 PT Claim 5; Page 5; 8pp; German.
 PS This invention describes a novel fluorogenic conjugate (vector) for
 XX fluorescence labeling of specific targets within cells which comprises a
 CC membrane translocation component, a targeting component and a
 CC fluorophore. A kit comprising two such conjugates, in which the
 CC fluorophores form a donor-acceptor pair for fluorescence resonance energy
 CC transfer (FRET), can be used to measure interactions between two
 CC substances, especially in living cells, when the targeting component of
 CC one of the conjugates targets one of the substances and the targeting
 CC component of the other conjugate targets the other substance. The
 CC conjugate or kit can be used to detect a target in a biologically
 CC functional cell. When the conjugate has a fluorophore capable of acting
 CC as a donor or acceptor for FRET with a fluorescent protein, it can be
 CC used to measure intracellular levels of a fusion protein of the
 CC fluorescent protein by FRET if the targeting component targets a sequence
 CC of the fusion protein, especially in a screening assay in which the gene
 CC encoding the fluorescent protein is linked to another coding sequence in
 CC an expression vector, the vector is used to transform a cell culture, the
 CC cell culture is incubated with the fluorogenic conjugate, and the cells
 CC are separated, preferably by means of a cell sorter or by FRET
 CC microscopy, especially where the fluorescent protein is green fluorescent
 CC protein (GFP) and the fluorophore is BODIPY, fluorescein, Oregon green,
 CC Rhodol green, rhodamine, Texas red, Cy2, Cy3, Cy5, Alexa, Marina blue,
 CC Pacific blue or AMCA. The cells are especially bacterial, yeast, insect,
 CC amphibian or mammalian cells. This sequence represents a fragment of the
 CC antennapedia homeodomain protein which is used to illustrate the method
 CC of the invention
 XX SQ Sequence 16 AA;
 Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 1 RQIKIWQNRRMKWKK 16
 |||||
 RESULT 21
 AAY51167
 ID AAY51167 standard; protein; 16 AA.
 XX AC AAY51167;
 XX 31-MAR-2000 (first entry)
 DT Drosophila sp. derived peptide fragment.
 XX TAT protein; nuclear transport; cell adhesion; gene therapy.
 KW Drosophila sp.
 XX WO9966061-A1.
 XX PN

XX 23-DEC-1999.
 XX 18-JUN-1999; 99WO-JP003272.
 XX 18-JUN-1998; 98JP-00189845.
 XX (DNAV-) Dनावे RES INC.
 XX Akuta T, Yokoi H, Okuyama H, Takeda K, Hasegawa M, Nakanishi M;
 XX WPI; 2000-097750/08.
 XX Phase for gene therapy carrying in its head part a bifunctional nuclear
 XX transport/cell adhesion protein.
 XX Disclosure; Page 38; 42pp; Japanese.
 XX This invention describes a novel phage whose head part contains a
 XX bifunctional protein with nuclear transport and cell adhesion activities.
 XX The phage is used as a vector for transport of foreign genes into a cell
 XX nucleus for gene therapy. This sequence represents a peptide fragment
 XX which is used in the method of the invention
 XX Sequence 16 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWQNRRMKWKK 16
 |||||
 DB 1 RQIKIWQNRRMKWKK 16
 |||||

RESULT 22
 AAB10343
 ID AAB10343 standard; protein; 16 AA.
 XX AC AAB10343;
 XX 24-NOV-2000 (first entry)
 XX Peptide AB fragment.
 XX Antiproliferative; nuclear localization sequence; retinoblastoma protein;
 XX antitumor; antiproliferative; antiatherosclerotic; treatment; diagnosis;
 XX benign tumor; malignant tumor; breast; atherosclerosis; psoriasis;
 XX chemotherapy.
 XX Synthetic.
 XX OS
 XX DE19859486-A1.
 XX 06-JUL-2000.
 XX 22-DEC-1998; 98DE-01059486.
 XX 22-DEC-1998; 98DE-01059486.
 XX (RADU/) RADULESCU R T.
 XX Radulescu RT;
 XX WPI; 2000-466905/41.
 XX New peptide with high antiproliferative activity, useful for treating
 XX cancer or psoriasis, comprises a nuclear localization sequence and a
 XX fragment of the retinoblastoma protein.
 XX Claim 1c; Page 7; 8pp; German.
 XX This invention describes novel peptides (I) with high antiproliferative

CC activity which comprise a nuclear localization sequence and a fragment of
 CC the retinoblastoma protein. The products of the invention have antitumor,
 CC antiproliferative, antiproliferative and antiatherosclerotic activity. The
 CC peptides are used to treat and/or diagnose diseases that involve
 CC increased proliferation or hyperproliferation of cells, specifically
 CC benign or malignant tumors (especially of the breast), atherosclerosis
 CC and psoriasis. (I) are more active against proliferation than similar
 CC compounds described in DE19653445, so doses and treatment times may be
 CC reduced, thus also reducing the cost. They destroy almost all types of
 CC tumor cells, especially those in which the retinoblastoma gene or protein
 CC is defective, but have relatively little effect on normal cells, which is
 CC in contrast to conventional chemotherapeutic agents. This sequence
 CC represents a peptide, AB which is used to describe the method of the
 CC invention
 XX Sequence 16 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKIWQNRRMKWKK 16
 |||||
 DB 1 RQIKIWQNRRMKWKK 16
 |||||

RESULT 23
 AAB19251
 ID AAB19251 standard; peptide; 16 AA.
 XX AC AAB19251;
 XX 19-FEB-2001 (first entry)
 XX Fragment of the Antennapedia protein from Drosophila.
 XX Antennapedia protein; translocating protein; cellular process;
 XX protein delivery.
 XX Drosophila sp.
 XX WO200058488-A2.
 XX 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US008571.
 XX 31-MAR-1999; 99US-0127467P.
 XX (INVI-) INVITROGEN CORP.
 XX Dalby B, Bennett RP;
 XX WPI; 2000-611716/58.
 XX Modulating a cellular process by contacting a cell in culture with a cell
 XX process modifying molecule attached to a translocating polypeptide,
 XX useful for modulating expression of a target gene product.
 XX Disclosure; Page 6; 59pp; English.
 XX The present sequence represents a fragment of the Antennapedia protein
 XX (amino acids 43-58) from Drosophila. The fragment is used as a
 XX translocating protein in the course of the invention. The specification
 XX describes a method for modulating a cellular process and for delivery of
 XX functional protein sequences. The method comprises contacting a cell in
 XX culture under suitable conditions with a cell process modifying molecule
 XX attached to a translocating polypeptide, where molecule is translocated
 XX into the cell and interacts specifically with a responsive target site.
 XX The method is useful for modulating a cellular process, such as
 XX modulating expression of a target gene product, of a cell in culture
 XX Sequence 16 AA;
 SQ

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
 |||||
 DB 1 RQIKWIFQNRMMKWK 16

RESULT 24
 AAY93178
 ID AAY93178 standard; peptide; 16 AA.
 XX
 AC
 XX
 XX
 DT 06-DEC-2000 (first entry)
 XX
 DE Protegrin-like peptide antibiotic No 2.
 XX
 KW Cytostatic; neurodegenerative; antidepressant; analgesic; antimicrobial;
 KW blood-brain barrier; diagnostic; central nervous system; protegrin;
 KW Antennapedia; tachyplesin; peptide antibiotic; Alzheimer's disease;
 KW cancer; Parkinson's disease; depression; pain; meningitis.
 XX
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "linked to doxorubicin via a succinimoyl
 FT maleimido-propionate-3-mercaptopropionate linker"
 FT
 XX WO200032236-A1.
 XX
 XX
 PD 08-JUN-2000.
 XX
 PF 26-NOV-1999; 99WO-FR002938.
 XX
 PR 30-NOV-1998; 98FR-00015074.
 XX
 XX (SYNT-) SYNT:EM SA.
 PA
 XX Clair P, Kaczorek M, Tamsamani J;
 PI WPI; 2000-422871/36.
 XX
 XX Use of linear peptides as vectors for active ingredients, useful for
 PT diagnosis and treatment of central nervous system diseases, can transport
 PT agents passively across the blood-brain barrier.
 XX
 PS Example I; Page 13; 54pp; French.
 XX
 CC The invention relates to the use of linear peptides, coupled to an active
 CC agent, to prepare a composition able to cross the blood-brain barrier
 CC for diagnosis or treatment of disorders localised in the central nervous
 CC system. The linear peptide preferably has the formula: (a) X1-X16; (b);
 CC EXBXXXXBBBXXXXXB; or (c) EXBXXXXBBBXXXXBBB, where: each of X1-X16 are
 CC amino acids (aa), of which 6-10 of them are hydrophobic and X6 must be
 CC Trp; each B is aa containing a side chain that includes a basic group;
 CC and each X is an aliphatic or aromatic aa. The linear peptide may be
 CC retro forms of (a)-(c) containing D- and/or L-form aa, or a fragment
 CC containing at least 5, preferably at least 7 consecutive aa from (a)-(c).
 CC Peptides able to cross the BBB include protegrins, Antennapedia,
 CC tachyplesins, transportan, etc. Of these several families have cytolytic
 CC effects and are termed peptide antibiotics. They fall into 3 main
 CC categories based on their structure: (i) peptides with alpha-helices,
 CC e.g. cecropins and maganins; (ii) peptides with disulphide bond-linked
 CC beta-sheets, e.g. protegrin, tachyplesins, defensins; (iii) peptides with
 CC no major structure but containing bends due to the presence of Pro
 CC residues, e.g. bactericins and PR39. The peptides of the invention fall
 CC into the peptide antibiotic categories defined above: (a)-peptides are
 CC based on the Antennapedia family peptides; (b)-peptides are based on
 CC protegrins; and (c)-peptides are based on tachyplesins. This sequence

CC represents a synthetic linear peptide designed on peptides able to cross
 CC the BBB and is conjugated to a doxorubicin molecule by a succinimoyl
 CC maleimido-propionate-3-mercaptopropionate linker. Conjugates of the
 CC linear peptides and the active agent are particularly used to treat,
 CC prevent or diagnose brain cancer, Alzheimer's or Parkinson's diseases,
 CC depression, pain and meningitis, but also for studying drug behaviour in
 CC BBB models
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
 |||||
 DB 1 RQIKWIFQNRMMKWK 16

RESULT 25
 AAB35694
 ID AAB35694 standard; peptide; 16 AA.
 XX
 AC AAB35694;
 XX
 DT 16-FEB-2001 (first entry)
 XX
 DE Peptide associated with modified beta-catenin expression #27.
 XX
 KW Beta-catenin; cell differentiation; hair growth; cancer;
 KW Alzheimer's disease.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200063246-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 21-APR-2000; 2000WO-US010753.
 XX
 PR 21-APR-1999; 99US-00296089.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 XX Blaschuk OW, Byers S, Gour BJ;
 XX WPI; 2000-679589/66.
 XX
 XX Use of modulating agent comprising internalization moiety and a peptide,
 PT for modulating beta-catenin mediated gene transcription and cell
 PT differentiation, for treating cancer, and for inhibiting Alzheimer's
 PT disease.
 XX
 XX Claim 6; Page 26; 47pp; English.
 PS
 CC The present invention relates to a method for modulating beta-catenin
 CC mediated gene transcription in a cell. The method involves contacting a
 CC cell with a modulating agent comprising an internalization moiety and a
 CC peptide comprising a sequence LXXLL or peptide analogue of LXXLL. The
 CC method is useful for modulating beta-catenin mediated gene transcription,
 CC cell differentiation, hair growth, and retinoic acid activity, for
 CC treating cancer, and for inhibiting the development of Alzheimer's
 CC disease
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
 |||||
 DB 1 RQIKWIFQNRMMKWK 16

Search completed: May 24, 2004, 17:22:19
Job time : 47.8108 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:18:06 ; Search time 9.2973 Seconds
(without alignments)
165.539 Million cell updates/sec

Title: US-09-977-349-1
Perfect score: 92
Sequence: 1 RQIKWFWQRNMKWK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 78.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	33	S57235	antennapedia prote
2	92	100.0	42	2	homeotic protein H
3	92	100.0	45	2	homeotic protein D
4	92	100.0	48	2	homeobox protein -
5	92	100.0	66	2	homeotic protein H
6	92	100.0	66	2	homeotic protein H
7	92	100.0	71	2	homeotic protein 3
8	92	100.0	71	2	homeotic protein H
9	92	100.0	74	2	homeo box protein
10	92	100.0	75	2	homeo box protein
11	92	100.0	75	2	homeotic protein S
12	92	100.0	75	2	homeotic protein R
13	92	100.0	78	2	homeo box protein
14	92	100.0	81	2	homeotic protein H
15	92	100.0	81	2	homeotic protein H
16	92	100.0	82	2	homeotic protein H
17	92	100.0	83	2	homeotic protein H
18	92	100.0	83	2	homeotic protein H
19	92	100.0	86	2	homeotic protein Z
20	92	100.0	86	2	homeotic protein H
21	92	100.0	86	2	homeotic protein H
22	92	100.0	87	2	homeotic protein H
23	92	100.0	88	2	homeotic protein Z
24	92	100.0	96	2	homeotic protein H
25	92	100.0	96	2	homeotic protein H
26	92	100.0	97	2	homeotic protein m
27	92	100.0	97	2	homeotic protein H
28	92	100.0	103	2	homeotic protein H
29	92	100.0	105	2	homeotic protein H

30	92	100.0	105	2	A27471	homeotic protein R
31	92	100.0	106	2	S36448	homeotic protein s
32	92	100.0	107	2	B61045	homeotic protein T
33	92	100.0	113	2	T10775	homeobox protein -
34	92	100.0	118	2	A24777	homeotic protein H
35	92	100.0	118	2	JT0273	homeotic protein H
36	92	100.0	118	2	B24777	homeotic protein M
37	92	100.0	119	2	A03314	homeotic protein m
38	92	100.0	138	2	S20087	homeotic protein b
39	92	100.0	148	2	PC4071	homeobox A5 protei
40	92	100.0	153	1	WJHU3C	homeotic protein H
41	92	100.0	153	1	WJMSX6	homeotic protein H
42	92	100.0	158	2	A27348	homeotic protein H
43	92	100.0	209	2	A43553	homeotic protein H
44	92	100.0	217	1	WJHU2C	homeotic protein H
45	92	100.0	217	1	WJMSX2	homeotic protein H
46	92	100.0	220	2	A37371	Hox 2 protein type
47	92	100.0	220	2	S01063	homeotic protein H
48	92	100.0	222	2	S20029	homeotic protein H
49	92	100.0	224	2	S26400	homeotic protein H
50	92	100.0	224	2	A31324	homeotic protein H
51	92	100.0	227	2	S50067	homeotic protein H
52	92	100.0	228	2	S32563	homeotic protein H
53	92	100.0	229	2	A28329	homeotic protein H
54	92	100.0	230	2	S00592	homeotic protein H
55	92	100.0	232	1	A25108	homeotic protein H
56	92	100.0	232	2	S48125	homeotic protein H
57	92	100.0	233	2	I65197	hox1.3 protein - x
58	92	100.0	234	1	S02014	homeotic protein H
59	92	100.0	234	2	S00992	homeotic protein H
60	92	100.0	235	2	A56568	homeotic protein H
61	92	100.0	235	2	S72429	homeotic protein H
62	92	100.0	236	1	S09256	homeotic protein H
63	92	100.0	242	1	A39164	homeotic protein H
64	92	100.0	245	1	S10092	homeotic protein H
65	92	100.0	246	2	T46446	hypothetical prote
66	92	100.0	250	1	A36170	homeotic protein H
67	92	100.0	250	1	A31757	homeotic protein H
68	92	100.0	251	1	B60492	homeotic protein H
69	92	100.0	255	1	WJHU4B	homeotic protein H
70	92	100.0	264	1	WJHU3E	homeotic protein H
71	92	100.0	264	1	S35219	homeotic protein H
72	92	100.0	269	1	A43551	homeotic protein H
73	92	100.0	269	1	A45578	homeotic protein H
74	92	100.0	270	1	WJHU1C	homeotic protein H
75	92	100.0	270	1	WJMS13	homeotic protein H
76	92	100.0	275	1	WJZFX2	homeotic protein H
77	92	100.0	295	2	S58850	homeotic protein A
78	92	100.0	378	2	A25399	homeotic protein A
79	92	100.0	394	2	S26492	homeotic protein A
80	92	100.0	413	1	S03631	homeotic protein S
81	92	100.0	590	2	A26638	homeotic protein D
82	91	98.9	60	2	PC2399	antennapedia-like
83	91	98.9	60	2	PC2400	antennapedia-like
84	91	98.9	60	2	I51437	homeotic protein H
85	91	98.9	62	2	B32391	homeotic protein H
86	91	98.9	66	2	S15534	homeotic protein R
87	91	98.9	67	2	B27471	homeotic protein m
88	91	98.9	69	2	S13785	homeotic protein H
89	91	98.9	70	2	B37042	homeotic protein H
90	91	98.9	81	2	A25180	homeotic protein H
91	91	98.9	85	2	I65198	homeotic protein H
92	91	98.9	85	2	A25472	homeotic protein H
93	91	98.9	95	2	B32830	homeotic protein H
94	91	98.9	96	2	A55278	homeotic protein H
95	91	98.9	99	2	A41605	homeotic protein H
96	91	98.9	104	2	D43559	homeotic protein R
97	91	98.9	114	2	A43559	homeotic protein R
98	91	98.9	155	2	S25846	homeotic protein H
99	91	98.9	188	2	I50145	homeotic protein H
100	91	98.9	242	1	WJMSX3	homeotic protein H

ALIGNMENTS

RESULT 1

S57235
antennapedia protein (clone p1105) - fruit fly (*Drosophila pseudoobscura*) (fragment)
C:Species: *Drosophila pseudoobscura*
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
C:Accession: S57235
R:Randazzo, F.M.; Seeger, M.A.; Huss, C.A.; Sweeney, M.A.; Cecil, J.K.; Kaufman, T.C.
Genetics 133, 319-330, 1993
A:Title: Structural changes in the antennapedia complex of *Drosophila pseudoobscura*.
A:Reference number: S57224
A:Accession: S57235
A:Molecule type: DNA
A:Residues: 1-33 <PAN>
A:Cross-references: EMBL:X77711
C:Genetics:
A:Gene: FlyBase:Antp
A:Cross-references: FlyBase:FBgn0012693
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-22/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
|||
DB 7 RQIKWQNRRMKWKK 22

RESULT 2

I65241
homeotic protein Hox-A - rat (fragment)
C:Species: *Rattus norvegicus* (Norway rat)
C>Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 15-Oct-1999
C:Accession: I65241
R:Sakoyama, Y.; Mizuta, I.; Ogasawara, N.; Yoshikawa, H.
Biochem. Genet. 32, 351-360, 1994
A:Title: Cloning of rat homeobox genes.
A:Reference number: I52340; MUID:95217128; PMID:7702549
A:Accession: I65241
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-42 <RES>
A:Cross-references: GB:S76290; NID:G913077
C:Genetics:
A:Gene: Hox-A; Hox-1
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
|||
DB 25 RQIKWQNRRMKWKK 40

RESULT 3

PC1216
homeotic protein Dthb1 - planarian (*Dugesia tigrina*) (fragment)
C:Species: *Dugesia tigrina*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Oct-1997
C:Accession: PC1216
R:Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlich
Gene 121, 337-342, 1992
A:Title: Homeoboxes in flatworms.
A:Reference number: JCI386; MUID:93077050; PMID:1359988

A:Accession: PC1216
A:Molecule type: DNA
A:Residues: 1-45 <OL1>
A:Cross-references: EMBL:X65822
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-45/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
|||
DB 30 RQIKWQNRRMKWKK 45

RESULT 4

I51439
homeobox protein - African clawed frog (fragment)
C:Species: *Xenopus laevis* (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 15-Oct-1999
C:Accession: I51439
R:Leeroy, P.; DeRobertis, E.M.
Dev. Dyn. 194, 21-32, 1992
A:Title: Effects of lithium chloride and retinoic acid on the expression of genes from t
A:Reference number: I51439; MUID:93043517; PMID:1384809
A:Accession: I51439
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-48 <LER>
A:Cross-references: GB:M91587; NID:G214257; PIDN:AAA49750.1; PID:G214258
C:Genetics:
A:Gene: Hox2.2
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-27/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
|||
DB 12 RQIKWQNRRMKWKK 27

RESULT 5

S15536
homeotic protein Hox A7 - human (fragment)
N:Alternate names: homeotic protein Hox 1A
C:Species: *Homo sapiens* (man)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
C:Accession: S15536
R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto
Genome 31, 745-756, 1989
A:Title: Organization of human class I homeobox genes.
A:Reference number: S15036; MUID:90215256; PMID:2576652
A:Accession: S15536
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-66 <BON>
C:Genetics:
A:Gene: GDB:HoxA7
A:Cross-references: GDB:120647; OMIM:142950
A:Map position: 7p15.3-7p15.3
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:2-58/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 43 RQIKIWQNRRMKWKK 58

RESULT 6

SI5538
 homeotic protein Hox A6 - human (fragment)
 N;Alternate names: homeotic protein Hox 1B
 C;Species: Homo sapiens (man)
 C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
 C;Accession: SI5538
 R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Stoc
 Genome 31, 745-756, 1989
 A;Title: Organization of human class I homeobox genes.
 A;Reference number: SI5036; MUID:90215256; PMID:2576652
 A;Accession: SI5538
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-66 <BON>
 C;Genetics:
 A;Gene: GDB:HOXA6
 A;Cross-references: GDB:120648; OMIM:142951
 A;Map position: 7p15.3-7p15.3
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;2-58/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 66;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0

Oy 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 43 RQIKIWQNRRMKWKK 58

RESULT 7

JC1161
 homeotic protein 3.4 - eastern newt (fragment)
 C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 23-May-1997
 C;Accession: JC1161
 R;Belleville, S.; Beauchemin, M.; Tremblay, M.; Noiseux, N.; Savard, P.
 Gene 114, 179-186, 1992
 A;Title: Homeobox-containing genes in the newt are organized in clusters similar to other
 A;Reference number: JC1161; MUID:92290273; PMID:1351019
 A;Accession: JC1161
 A;Molecule type: DNA
 A;Residues: 1-71 <BEL>
 A;Cross-references: GB:M84001
 C;Genetics:
 A;Gene: NVHBox-3.4
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;5-61/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 71;
 Best Local Similarity 100.0%; Pred. No. 3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 46 RQIKIWQNRRMKWKK 61

RESULT 8

A60084
 homeotic protein Hox 3.4 - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 07-May-1999
 C;Accession: A60084

R;Gaunt, S.J.; Coletta, P.L.; Pravtcheva, D.; Sharpe, P.T.

Development 109, 329-339, 1990
 A;Title: Mouse Hox-3.4: homeobox sequence and embryonic expression patterns compared wit
 A;Reference number: A60084; MUID:90382249; PMID:1976089
 A;Accession: A60084

A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-71 <GAU>

C;Genetics:
 A;Map position: 15
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;5-61/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 71;
 Best Local Similarity 100.0%; Pred. No. 3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 46 RQIKIWQNRRMKWKK 61

RESULT 9

D34510
 homeotic protein H90 - honeybee (fragment)
 C;Species: Apis mellifera (honeybee)
 C;Date: 22-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 24-Sep-1999
 C;Accession: D34510
 R;Walldorf, U.; Fleig, R.; Gehring, W.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 9971-9975, 1989
 A;Title: Comparison of homeobox-containing genes of the honeybee and Drosophila.
 A;Reference number: A34510; MUID:90099384; PMID:2574865
 A;Accession: D34510
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-74 <WAL>
 A;Cross-references: GB:M29493; NID:G155675; PIDN:AAA27728.1; PID:G155676
 A;Note: the authors mistranslated the codons for residues 68-74
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;9-65/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 74;
 Best Local Similarity 100.0%; Pred. No. 3.1e-07; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0

Oy 1 RQIKIWQNRRMKWKK 16
 |||||
 Db 50 RQIKIWQNRRMKWKK 65

RESULT 10

IS1341
 homeo box protein - Atlantic salmon (fragment)
 C;Species: Salmo salar (Atlantic salmon)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
 C;Accession: IS1341
 R;Fjose, A.; Molven, A.; Eiken, H.G.
 Gene 62, 141-152, 1988
 A;Title: Molecular cloning and characterization of homeobox-containing genes from Atlant:
 A;Reference number: IS1341; MUID:88226009; PMID:2897318
 A;Accession: IS1341
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-75 <FJO>
 A;Cross-references: GB:MI8903; NID:G213797; PIDN:AAA49559.1; PID:G213798
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;2-58/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 75;
 Best Local Similarity 100.0%; Pred. No. 3.1e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||
 Db 43 RQIKWFWQRNRMKWKK 58

RESULT 11

S58852 homeotic protein Scr homolog - Junonia coenia (fragment)
 N; Alternate names: sex combs reduced homeodomain protein
 C; Species: Junonia coenia
 C; Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 24-Sep-1999
 C; Accession: S58852
 R; Warren, R.W.; Nagy, L.; Selegue, J.; Gates, J.; Carroll, S.
 Nature 372, 458-461, 1994
 A; Title: Evolution of homeotic gene regulation and function in flies and butterflies.
 A; Reference number: S58850; MUID:95075456; PMID:7840822
 A; Accession: S58852
 A; Status: nucleic acid sequence not shown; translation not shown
 A; Molecule type: mRNA
 A; Residues: 1-75 <VAR>

A; Cross-references: EMBL:L42136; NID:G806495; PIDN:AAA68462.1; PID:G806496
 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F; 1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 75;
 Best Local Similarity 100.0%; Pred. No. 3.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||
 Db 25 RQIKWFWQRNRMKWKK 40

RESULT 12

C43559 homeotic protein R3 - rat (fragment)
 C; Species: Rattus norvegicus (Norway rat)
 C; Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 24-Sep-1999
 C; Accession: C43559
 R; Palzon, M.; Chung, S.Y.
 Development 103, 601-610, 1988
 A; Title: The expression of rat homeobox-containing genes is developmentally regulated and
 A; Reference number: A43559; MUID:89231502; PMID:2907739
 A; Accession: C43559
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-76 <PAL>

A; Cross-references: GB:M37567; NID:G204634; PIDN:AAA41343.1; PID:G204635
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F; 20-76/Domain: homeobox homology (fragment) <HOX>

Query Match 100.0%; Score 92; DB 2; Length 76;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||
 Db 53 RQIKWFWQRNRMKWKK 68

RESULT 13

I51342 homeo box protein - Atlantic salmon (fragment)
 C; Species: Salmo salar (Atlantic salmon)
 C; Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
 C; Accession: I51342
 R; Fjose, A.; Molven, A.; Eiken, H.G.
 Gene 62, 141-152, 1988

A; Title: Molecular cloning and characterization of homeobox-containing genes from Atlant
 A; Reference number: I51341; MUID:88226009; PMID:2897318
 A; Accession: I51342
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 1-78 <EJO>

A; Cross-references: GB:M18904; NID:G213799; PIDN:AAA49560.1; PID:G213800
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F; 2-58/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 78;
 Best Local Similarity 100.0%; Pred. No. 3.2e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||
 Db 43 RQIKWFWQRNRMKWKK 58

RESULT 14

S47605 homeotic protein Hox-7 - Florida lancelet (fragment)
 C; Species: Branchiostoma floridae (Florida lancelet)
 C; Date: 01-Feb-1995 #sequence_revision 26-May-1995 #text_change 24-Sep-1999
 C; Accession: S47605
 R; Garcia-Fernandez, J.; Holland, P.W.H.
 Nature 370, 563-566, 1994

A; Title: Archetypal organization of the amphioxus Hox gene cluster.
 A; Reference number: S47599; MUID:94329179; PMID:7914353
 A; Accession: S47605
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-81 <GAR>

A; Cross-references: EMBL:Z35147; NID:G520617; PIDN:CAA84519.1; PID:G520618
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F; 4-60/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||
 Db 45 RQIKWFWQRNRMKWKK 60

RESULT 15

B29585 homeotic protein Hox 2.2 precursor - mouse (fragment)
 C; Species: Mus musculus (house mouse)
 C; Date: 15-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 17-Oct-1997
 C; Accession: B29585
 R; Lönai, P.; Arman, E.; Czosnek, H.; Ruddle, F.H.; Blatt, C.
 DNA 6, 409-418, 1987

A; Title: New murine homeoboxes: structure, chromosomal assignment, and differential expr
 A; Reference number: A29585; MUID:88054465; PMID:2890503
 A; Accession: B29585
 A; Molecule type: DNA
 A; Residues: 1-81 <LON>

A; Cross-references: GB:M18167
 A; Note: the authors translated the codon CAG for residue 69 as Glu
 C; Superfamily: unassigned homeobox proteins; homeobox homology
 C; Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F; 4-60/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRNRMKWKK 16
 |||||

Db 45 RQIKWIFQNRRMKWK 60

RESULT 16

S08302
 homeotic protein Hox 2.1 - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C>Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
 C:Accession: S08302
 R:Wedden, S.E.; Pang, K.; Eichele, G.
 Development 105, 639-650, 1989
 A:Title: Expression pattern of homeobox-containing genes during chick embryogenesis.
 A:Reference number: S08302; MUID:90126373; PMID:2575515
 A:Accession: S08302
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-82 <WED>
 A:Cross-references: EMBL:X16846; NID:962905; PIDN:CAA34743.1; PID:gl334633
 A:Note: the authors translated the codon GGA for residue 4 as Arg
 C:Superfamily: homeotic protein Hox A5; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;8-64/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 82;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16

Db 49 RQIKWIFQNRRMKWK 64

RESULT 17

S47603
 homeotic protein Hox-5 - Florida lancelet (fragment)
 C:Species: Branchiostoma floridae (Florida lancelet)
 C>Date: 01-Feb-1995 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
 C:Accession: S47603
 R:Garcia-Fernandez, J.; Holland, P.W.H.
 Nature 370, 563-566, 1994
 A:Title: Archetypal organization of the amphioxus Hox gene cluster.
 A:Reference number: S47599; MUID:94329179; PMID:7914353
 A:Accession: S47603
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-83 <GAR>
 A:Cross-references: EMBL:Z35145
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;7-63/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16

Db 48 RQIKWIFQNRRMKWK 63

RESULT 18

S50066
 homeotic protein Hox-5 - Florida lancelet (fragment)
 C:Species: Branchiostoma floridae (Florida lancelet)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
 C:Accession: S50066
 R:Holland, P.W.
 submitted to the EMBL Data Library, July 1994
 A:Reference number: S50066
 A:Accession: S50066
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-83 <HOL>

A:Cross-references: EMBL:Z35145; NID:9520613; PIDN:CAA84517.1; PID:g520614
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;7-63/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 3.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16

Db 48 RQIKWIFQNRRMKWK 63

RESULT 19

A34510
 homeotic protein H55 - honeybee (fragment)
 C:Species: Apis mellifera (honeybee)
 C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 24-Sep-1999
 C:Accession: A34510
 R:Waldorf, U.; Fleig, R.; Gehring, W.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 9971-9975, 1989
 A:Title: Comparison of homeobox-containing genes of the honeybee and Drosophila.
 A:Reference number: A34510; MUID:9009384; PMID:2574865
 A:Accession: A34510
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-86 <WAL>
 A:Cross-references: GB:M29488; NID:g155665; PIDN:AAA27723.1; PID:g155666
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;9-65/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16

Db 50 RQIKWIFQNRRMKWK 65

RESULT 20

JT0489
 homeotic protein ZF-13 - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 17-Oct-1997
 C:Accession: JT0489
 R:Njoelstad, P.R.; Molven, A.; Eiken, H.G.; Fjose, A.
 Gene 73, 33-46, 1988
 A:Title: Structure and neural expression of a zebrafish homeobox sequence.
 A:Reference number: JT0489; MUID:89211958; PMID:2468579
 A:Accession: JT0489
 A:Molecule type: DNA
 A:Residues: 1-86 <NJO>
 C:Genetics:
 A:Gene: ZF-13
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;11-67/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16

Db 52 RQIKWIFQNRRMKWK 67

RESULT 21

S08303
 homeotic protein Hox 2.2 - chicken (fragment)

C:Species: Gallus gallus (chicken)
 C>Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 17-Oct-1997
 C:Accession: S08303
 R:Wedden, S.E.; Pang, K.; Eichele, G.
 Development 105, 639-650, 1989
 A:Title: Expression pattern of homeobox-containing genes during chick embryogenesis.
 A:Reference number: S08302; MUID:90126373; PMID:2575515
 A:Accession: S08303
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-86 <WED>
 A:Cross-references: EMBL:X16847
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:10-66/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 51 RQIKWIFQNRMKWK 66

RESULT 22
 S00589
 homeotic protein Hox 5 - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 24-Sep-1999
 C:Accession: S00589
 R:Pritz, A.; de Robertis, E.M.
 Nucleic Acids Res. 16, 1453-1469, 1988
 A:Title: Xenopus homeobox-containing cDNAs expressed in early development.
 A:Reference number: S00589; MUID:88157707; PMID:2894634
 A:Accession: S00589
 A:Molecule type: mRNA
 A:Residues: 1-87 <FRI>
 A:Cross-references: EMBL:X07105; NID:964757; PIDN:CAA30126.1; PID:e12307; PID:gl334653
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:21-77/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 62 RQIKWIFQNRMKWK 77

RESULT 23
 A03317
 homeotic protein MM3 - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Sep-1999
 C:Accession: A03317
 R:Muller, M.M.; Carrasco, A.E.; DeRobertis, E.M.
 Cell 39, 157-162, 1984
 A:Title: A homeo-box-containing gene expressed during oogenesis in Xenopus.
 A:Reference number: A03317; MUID:85024899; PMID:6207937
 A:Accession: A03317
 A:Molecule type: DNA
 A:Residues: 1-88 <MUL>
 A:Cross-references: GB:X02616; NID:g214254; PIDN:AAA49749.1; PID:g214255
 C:Comment: This protein is expressed in oocytes.
 C:Genetics:
 A:Gene: MM3
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: alternative splicing; DNA binding; homeobox; nucleus; oocyte; transcription
 F:9-65/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16

Query Match 100.0%; Score 92; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 50 RQIKWIFQNRMKWK 65

RESULT 24
 S08639
 homeotic protein zf-61 - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
 C:Accession: S08639
 R:Njolstad, P.R.; Molven, A.; Apold, J.; Fjose, A.
 EMBO J. 9, 515-524, 1990
 A:Title: The zebrafish homeobox gene hox-2.2: transcription unit, potential regulatory r
 A:Reference number: S08639; MUID:90151628; PMID:1968004
 A:Accession: S08639
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-96 <NJO>
 A:Cross-references: EMBL:X17266; NID:g62538; PIDN:CAA35170.1; PID:e16657; PID:gl334622
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:18-64/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 49 RQIKWIFQNRMKWK 64

RESULT 25
 A05266
 homeotic protein Hox B6 - human (fragment)
 N:Alternate names: homeotic protein Hox 2B; homeotic protein Hu2
 C:Species: Homo sapiens (man)
 C>Date: 05-Jun-1987 #sequence_revision 30-Sep-1991 #text_change 17-Oct-1997
 C:Accession: A05266; S15537
 R:Levine, M.; Rubin, G.M.; Tjian, R.
 Cell 38, 667-673, 1984
 A:Title: Human DNA sequences homologous to a protein coding region conserved between hom
 A:Reference number: A05265; MUID:85024858; PMID:6091895
 A:Accession: A05266
 A:Molecule type: DNA
 A:Residues: 1-96 <LEV>
 A:Cross-references: EMBL:X02571
 A>Note: this reading frame extends between two stop codons and does not begin with a sta
 R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sto
 Genome 31, 745-756, 1989
 A:Title: Organization of human class I homeobox genes.
 A:Reference number: S15036; MUID:90215256; PMID:2576652
 A:Accession: S15537
 A:Molecule type: DNA
 A:Residues: 18-19, 'R', 21-83 <BON>
 C:Genetics:
 A:Gene: GDB:H0XB6
 A:Cross-references: GDB:120659; OMIM:142961
 A:Map position: 17q21.3-17q21.3
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:19-75/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 92; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16

Db 60 RQIKLWFONRMMKKK 75

Search completed: May 24, 2004, 17:26:10
Job time : 9.2973 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:17:30 ; Search time 5.83784 Seconds
(without alignments)
142.711 Million cell updates/sec

Title: US-09-977-349-1

Perfect score: 92
Sequence: 1 RQIKINFQRRMKWK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	48	1 HXB6_XENIA	P1256 xenopus lae
2	92	100.0	49	1 HXB5_SHEEP	Q28599 oviv aries
3	92	100.0	71	1 HXA7_HUMAN	Q28600 oviv aries
4	92	100.0	71	1 HXC5_NOTVI	P31262 notophthalm
5	92	100.0	74	1 HM90_APIPE	P15860 apis mellif
6	92	100.0	75	1 HXSA_SALSA	P09636 salmo salar
7	92	100.0	76	1 HXC4_RAT	P18865 rattus norv
8	92	100.0	78	1 HXA5_SALSA	P09637 salmo salar
9	92	100.0	80	1 HXA4_LINSA	P81192 lineus sang
10	92	100.0	81	1 HX5L_BRARE	P09013 brachydantio
11	92	100.0	82	1 HXB6_CHICK	P14838 gallus gall
12	92	100.0	84	1 HXB6_CHICK	P14839 gallus gall
13	92	100.0	86	1 SCR_APIPE	P15859 apis mellif
14	92	100.0	87	1 HXC5_XENIA	P09020 xenopus lae
15	92	100.0	93	1 HXB8_PIG	P09078 sus scrofa
16	92	100.0	96	1 HXC6_BRARE	P15862 brachydantio
17	92	100.0	105	1 HXA7_RAT	P09634 rattus norv
18	92	100.0	105	1 HXB4_BRARE	P22574 brachydantio
19	92	100.0	112	1 HXB7_RAT	P18864 rattus norv
20	92	100.0	148	1 HXA5_AMBME	P50208 ambystoma m
21	92	100.0	153	1 HXC6_SHEEP	P49925 oviv aries
22	92	100.0	208	1 HXA7_HETFR	Q9125 heterodontu
23	92	100.0	209	1 HXA7_XENIA	P09071 xenopus lae
24	92	100.0	217	1 HXB7_BOVIN	Q9125 heterodontu
25	92	100.0	217	1 HXB7_HUMAN	P09629 homo sapien
26	92	100.0	217	1 HXB7_MOUSE	P09024 mus musculu
27	92	100.0	220	1 H87A_XENIA	Q91771 xenopus lae
28	92	100.0	220	1 H87B_XENIA	P04476 xenopus lae
29	92	100.0	222	1 HXC5_HUMAN	Q00444 homo sapien
30	92	100.0	222	1 HXC5_MOUSE	P32043 mus musculu
31	92	100.0	224	1 HXB6_HUMAN	P17509 homo sapien
32	92	100.0	224	1 HXB6_MOUSE	P09023 mus musculu
33	92	100.0	225	1 HXA7_MORSA	Q9pwa4 morone saxa

34	92	100.0	228	1 HXB6_BRARE	P15861 brachydantio
35	92	100.0	229	1 HXA6_HETFR	Q91a24 heterodontu
36	92	100.0	229	1 HXA7_MOUSE	P02830 mus musculu
37	92	100.0	230	1 HXA7_HUMAN	P31268 homo sapien
38	92	100.0	230	1 HXB5_XENIA	P09019 xenopus lae
39	92	100.0	232	1 HXA6_MOUSE	P09092 mus musculu
40	92	100.0	232	1 HXB4_XENIA	P09070 xenopus lae
41	92	100.0	232	1 HXC5_BRARE	P09074 brachydantio
42	92	100.0	233	1 HXA5_RAT	P52949 rattus norv
43	92	100.0	233	1 HXA6_HUMAN	P31267 homo sapien
44	92	100.0	234	1 HXC6_NOTVI	P14858 notophthalm
45	92	100.0	234	1 HXC6_XENIA	P02832 xenopus lae
46	92	100.0	235	1 HXC6_HUMAN	P09630 homo sapien
47	92	100.0	235	1 HXC6_MOUSE	P10629 mus musculu
48	92	100.0	235	1 HXD4_CHICK	P17278 gallus gall
49	92	100.0	236	1 HXD4_BRARE	O57374 brachydantio
50	92	100.0	242	1 HXA7_COTJA	P24061 coturnix co
51	92	100.0	245	1 HXB4_CHICK	P14840 gallus gall
52	92	100.0	250	1 HXB4_MOUSE	P10284 mus musculu
53	92	100.0	250	1 HXD4_MOUSE	P10628 mus musculu
54	92	100.0	251	1 HXB4_FUGRU	O13074 fugu rubrip
55	92	100.0	251	1 HXB4_HUMAN	P17483 homo sapien
56	92	100.0	252	1 HXD5_HETFR	Q91a11 heterodontu
57	92	100.0	255	1 HXD4_HUMAN	P09016 homo sapien
58	92	100.0	261	1 HXC4_ORYIA	Q9pva4 cryzias lat
59	92	100.0	264	1 HXC4_HUMAN	P09017 homo sapien
60	92	100.0	264	1 HXC4_MOUSE	Q08624 mus musculu
61	92	100.0	269	1 HXB5_HUMAN	P09067 homo sapien
62	92	100.0	269	1 HXB5_MOUSE	P09079 mus musculu
63	92	100.0	270	1 HXA5_HUMAN	P20719 homo sapien
64	92	100.0	270	1 HXA5_MOUSE	P09021 mus musculu
65	92	100.0	275	1 HXA5_HETFR	Q91a23 heterodontu
66	92	100.0	275	1 HXB5_BRARE	P09014 brachydantio
67	92	100.0	281	1 HXA5_MORSA	Q9pwa3 morone saxa
68	92	100.0	378	1 HMAN_DROME	P02833 drosophila
69	92	100.0	394	1 HMAN_DROSU	Q24645 drosophila
70	92	100.0	415	1 SCR_DROME	P09077 drosophila
71	92	100.0	590	1 HMDF_DROME	P07548 drosophila
72	91	98.9	67	1 HXA4_RAT	P09635 rattus norv
73	91	98.9	92	1 HXB8_RAT	P18863 rattus norv
74	91	98.9	108	1 HXC8_RAT	P18866 rattus norv
75	91	98.9	188	1 HXD8_CHICK	P23459 gallus gall
76	91	98.9	240	1 HXD8_HETFR	Q91a12 heterodontu
77	91	98.9	242	1 HXC8_HUMAN	P31273 homo sapien
78	91	98.9	242	1 HXC8_MOUSE	P09025 mus musculu
79	91	98.9	243	1 HXB8_HUMAN	P17481 homo sapien
80	91	98.9	243	1 HXB8_MOUSE	P09632 mus musculu
81	91	98.9	247	1 HXA4_HETFR	Q91a22 heterodontu
82	91	98.9	248	1 HXA4_MORSA	Q9pwa2 morone saxa
83	91	98.9	285	1 HXA4_MOUSE	P06798 mus musculu
84	91	98.9	289	1 HXD8_MOUSE	P23463 mus musculu
85	91	98.9	290	1 HXD8_HUMAN	P13378 homo sapien
86	91	98.9	309	1 HXA4_CHICK	P17277 gallus gall
87	91	98.9	320	1 HXA4_HUMAN	Q00056 homo sapien
88	89	96.7	108	1 HMB3_TRIGR	P10178 tripneustes
89	87	94.6	60	1 HMA2_HETFR	P17138 helobdella
90	87	94.6	253	1 MOX1_MOUSE	P32442 mus musculu
91	87	94.6	254	1 MOX1_HUMAN	P50221 homo sapien
92	87	94.6	271	1 HM8_XENIA	P14837 xenopus lae
93	87	94.6	283	1 IPF1_HUMAN	P52945 h insulin p
94	87	94.6	283	1 IPF1_MESAU	P70118 mesocricetu
95	87	94.6	283	1 IPF1_RAT	P52947 rattus norv
96	87	94.6	284	1 IPF1_MOUSE	P52946 mus musculu
97	87	94.6	298	1 MOX2_XENIA	P39021 xenopus lae
98	87	94.6	303	1 MOX2_HUMAN	P50222 homo sapien
99	87	94.6	303	1 MOX2_MOUSE	P32443 mus musculu
100	87	94.6	303	1 MOX2_RAT	P39020 rattus norv

ALIGNMENTS

RESULT 1

```

HXB6 XENLA
ID HXB6 XENLA STANDARD; PRT; 48 AA.
AC P31256;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B6 (XlHox-2.2) (Fragment).
GN HXB6 OR XLHox-2.2
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93043517; PubMed=1384809;
RA Leroy P., de Robertis E.M.;
RT "Effects of lithium chloride and retinoic acid on the expression of
RL genes from the Xenopus laevis Hox 2 complex.";
RL Dev. Dyn. 194:21-32(1992).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M91587; AAA49750.1; -.
CC PIR; I51439; I51439.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC DNA_BIND <1 29 HOMEBOX.
CC SEQUENCE 48 AA; 5716 MW; BC39E36822EDDD2A CRC64;

Query Match 100.0%; Score 92; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
DB 12 RQIKWIFQNRMKWK 27

RESULT 2
HXA5 SHEEP
ID HXA5 SHEEP STANDARD; PRT; 49 AA.
AC Q28599;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (Fragment).
GN HXA5 OR HOXA-5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Caprinae; Ovis.

Query Match 100.0%; Score 92; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
DB 12 RQIKWIFQNRMKWK 27

RESULT 2
HXA5 SHEEP
ID HXA5 SHEEP STANDARD; PRT; 49 AA.
AC Q28599;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (Fragment).
GN HXA5 OR HOXA-5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Caprinae; Ovis.

```

```

OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Roche P.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC Also binds to its own promoter. Binds specifically to the motif:
CC 5'-CYNATATA[NG]Y-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U61978; AAB04754.1; -.
CC HSSP; P02833; IHOM.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC DNA_BIND <1 49 HOMEBOX.
CC SEQUENCE 49 AA; 6331 MW; 1EE702315E7C099B CRC64;

Query Match 100.0%; Score 92; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
DB 32 RQIKWIFQNRMKWK 47

RESULT 3
HXA7 SHEEP
ID HXA7 SHEEP STANDARD; PRT; 71 AA.
AC Q28600;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A7 (Fragment).
GN HXA7 OR HOXA-7.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Roche P.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```
Matches 16; Conservative 0; Mismatches 0; Index 0
```

QY 1 RQIKWQNRMRKWK 16
 DB 50 RQIKWQNRMRKWK 65

RESULT 6

HMSA SALSA STANDARD; PRT; 75 AA.
 AC P09636;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Homeobox protein S12-A (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=88226009; PubMed=2897318;
 RA Fjose A., Molven A., Eiken H.G.;
 RT "Molecular cloning and characterization of homeo-box-containing genes
 from Atlantic salmon";
 RL Gene 62:141-152(1988).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M18903; AAA49559.1; -;
 FJ; I51341; I51341.

DR HSSP; P02833; 9ANT.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.

FT NON_TER 1 1
 FT DNA_BIND 1 60 HOMEBOX.
 FT NON_TER 75 75

SQ SEQUENCE 75 AA; 9330 MW; FC02C3672F35475D CRC64;

Query Match 100.0%; Score 92; DB 1; Length 75;

Best Local Similarity 100.0%; Pred. No. 4.1e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRMRKWK 16

DB 43 RQIKWQNRMRKWK 58

RESULT 7

HXC4 RAT STANDARD; PRT; 76 AA.
 AC P18865;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-C4 (R3) (Fragment).
 OS HOXC4 OR HOXC-4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=89231502; PubMed=2907739;
 RA Falzon M., Chung S.Y.;
 RT "The expression of rat homeobox-containing genes is developmentally
 regulated and tissue specific";
 RL Development 103:601-610(1988).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M37567; AAA41343.1; -;
 DR FJ; C43559; C43559.

DR HSSP; P02833; 9ANT.

DR InterPro; IPR001827; Antennapedia.

DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR000047; HTH lambrpressor.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR PRINTS; PR00031; HTHREPRESSR.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS50071; HOMEBOX 2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;

FT NON_TER 1 1

FT DNA_BIND 11 70 HOMEBOX.

SQ SEQUENCE 76 AA; 9293 MW; 5235F65C0672385 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 76;

Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRMRKWK 16

DB 53 RQIKWQNRMRKWK 68

RESULT 8

HXA5 SALSA STANDARD; PRT; 78 AA.
 AC P09637;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-A5 (S12-B) (Fragment).
 OS HOXA5.
 OS Salmo salar (Atlantic salmon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.

OX NCBI_TaxID=8030;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88226009; PubMed=2897318;

RA Fjose A., Molven A., Eiken H.G.;

InterPro; IPR001356; Homeobox.
InterPro; IPR000047; HTH_lambdarepressr.
Pfam; PF00046; homeobox; 1.
PRINTS; PR00024; HOMEBOX.
PRINTS; PR00031; HTHREPRESSR.
ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEBOX_1; 1.
PROSITE; PS00071; HOMEBOX_2; 1.
PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation.
FT NON TER 1
FT DNA_BIND 21 80 HOMEBOX.
FT NON TER 80 80
FT SEQUENCE 80 AA; 9860 NW; F2CE1B01CB8042F1 CRC64;
Query Match 100.0%; Score 92; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWIFQNRMKWK 16
|||
DB 53 RQIKWIFQNRMKWK 68
|||
RESULT 10
HX5L BRARE
ID ID_HX5L BRARE STANDARD; PRT; 81 AA.
AC P09013;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B5 like (ZF-54) (Fragment).
OS HOXB5B OR ZF54 OR ZF-54.
GN Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
ON NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=89016617; PubMed=2902580;
RA Njolstad P.R., Molven A., Hordvik I., Apold J., Fjose A.;
RT "Primary structure, developmentally regulated expression and
RT potential duplication of the zebrafish homeobox gene ZF-21.";
RL Nucleic Acids Res. 16:9097-9113(1988).
CC -1- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X12803; CAA31291.1; --
CC DR HSSP; P02833; ISAN.
CC ZFIN; ZDB-GENE-000823-6; hoxb5b.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON TER 1
FT DNA_BIND 21 80 HOMEBOX.
FT NON TER 80 80
FT SEQUENCE 80 AA; 9860 NW; F2CE1B01CB8042F1 CRC64;
Query Match 100.0%; Score 92; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 4.e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWIFQNRMKWK 16
|||
DB 53 RQIKWIFQNRMKWK 68
|||

Qy	Rel.	Created
1 ROIKIWONREMKWK 16	P19859;	AC P19859; 01-APR-1990 (Rel. 14, Created)
1	ROIKIWONREMKWK	16


```

DT 01-APR-1990 (Rel. 14, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Homeobox protein H55 (Fragment).
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9009384; PubMed=2574865;
RA Walldorf U., Fleig R., Gehring W.J.;
RT "Comparison of homeobox-containing genes of the honeybee and
RT Drosophila."
RL Proc. Natl. Acad. Sci. U.S.A. 86:9971-9975(1989).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ANT-P HOMEOBOX FAMILY. STRONGEST, TO SCR
CC OF DROSOPHILA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M29488; AA27723.1; -.
CC PIR; A34510; A34510.
CC HSSP; P02833; 1SAN.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS50071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT NON_TER 1
FT DNA_BIND 8 67 HOMEBOX.
FT NON_TER 86
FT NON_TER 86
SQ SEQUENCE 86 AA; 10713 MW; 2A49AB857C138AB8 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
DB 50 RQIKWQNRRMKWKK 65

RESULT 14
HXCS XENLA STANDARD; PRT; 87 AA.
AC P09020;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C5 (Xlhb-5) (Fragment).
GN HXC5 OR XLHBX5.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=88157707; PubMed=2894634;
RA Fritz A., de Robertis E.M.;
RT "Xenopus homeobox-containing cDNAs expressed in early development.";
RL Nucleic Acids Res. 16:1453-1469(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EXCLUSIVELY IN EARLY EMBRYOS.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X07105; CAA30126.1; -.
CC PIR; S00589; S00589.
CC HSSP; P02833; 9ANT.
CC TRANSFAC; T03765; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS50071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT NON_TER 1
FT SITE 5 10 ANT-P-TYPE HEXAPEPTIDE.
FT DNA_BIND 20 79 HOMEBOX.
FT NON_TER 1
FT SITE 1 1
SQ SEQUENCE 87 AA; 11058 MW; E67939E334E2BA43 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 4.8e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWQNRRMKWKK 16
DB 62 RQIKWQNRRMKWKK 77

RESULT 15
HXB8 PIG STANDARD; PRT; 93 AA.
AC P03078;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B8 (Hox-2.4) (Fragment).
GN HXB8 OR HOX-2.4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89057478; PubMed=2904133;
RA Miller J.R., Gaunt S.J., Sharpe P.T.;
RT "Pig Hox-2.4 has accumulated a frameshift mutation relative to mouse
RT Hox-2.4."
RL Nucleic Acids Res. 16:10364-10364(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.

```

CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06668; CAB57825.1; ALT_SEQ.
CC HSP; P02834; LB81.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambrpressor.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1
CC DNA_BIND 1
CC SEQUENCE 93 AA; 10872 MW; 7841DD6D17634ERC CRC64;
CC
CC Query Match 100.0%; Score 92; DB 1; Length 93;
CC Best Local Similarity 100.0%; Pred. No. 5.1e-08;
CC Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 RQIKWIFQNRMKWK 16
CC DB 49 RQIKWIFQNRMKWK 64
CC
CC RESULT 16
CC HXC6 BRARE
CC ID HXC6 BRARE STANDARD; PRT; 96 AA.
CC AC F15862;
CC DT 01-APR-1990 (Rel. 14, Created)
CC DT 01-APR-1990 (Rel. 14, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Homeobox protein Hox-C6 (ZF-61) (Fragment).
CC GN HXC6 OR HXC6A OR HXC-C6 OR ZF-61.
CC OS Brachydanio rerio (Zebrafish) (Danio rerio).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
CC OC Cyprinidae; Danio.
CC NCBI_TaxID=7955;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=Embryo;
CC MEDLINE=90151628; PubMed=1968004;
CC NJolstad P.R., Molven A., Apold J., Fjose A.;
CC "The zebrafish homeobox gene hox-2.2: transcription unit, potential
CC regulatory regions and in situ localization of transcripts.";
CC EMBO J. 9:515-524(1990).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC EMBL; X17266; CAA35170.1; -.
CC PIR; S08639; S08639.
CC HSP; P02833; 9ANT.
CC ZFIN; ZDB-GENE-990415-113; hoxc6a.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1
CC DNA_BIND 1
CC SEQUENCE 96 AA; 11638 MW; F1ED7AFAA3B640C0 CRC64;
CC
CC Query Match 100.0%; Score 92; DB 1; Length 96;
CC Best Local Similarity 100.0%; Pred. No. 5.3e-08;
CC Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 RQIKWIFQNRMKWK 16
CC DB 49 RQIKWIFQNRMKWK 64
CC
CC RESULT 17
CC HXA7 RAT
CC ID HXA7 RAT STANDARD; PRT; 105 AA.
CC AC P09634;
CC DT 01-MAR-1989 (Rel. 10, Created)
CC DT 01-MAR-1989 (Rel. 10, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Homeobox protein Hox-A7 (Hox-1.1) (R5) (Fragment).
CC GN HXA7 OR HOXA-7.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=Sprague-Dawley;
CC MEDLINE=87277429; PubMed=2886401;
CC Falzon M., Sanderson N., Chung S.Y.;
CC "Cloning and expression of rat homeo-box-containing sequences.";
CC Gene 54:23-32(1987).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M16807; -; NOT_ANNOTATED_CDS.
CC PIR; A27471; A27471.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T01707; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 Transcription regulation.
 FT NON_TER 1 1
 FT DNA_BIND 5 64 HOMEBOX.
 FT DOMAIN 91 105 GLU-RICH (ACIDIC).
 SQ SEQUENCE 105 AA; 12552 MW; 106CIDF938F2864B CRC64;

Query Match 100.0%; Score 92; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRKWK 16
 DB 47 RQIKWIFQNRMRKWK 62

RESULT 18

HXB4 BRARE STANDARD; PRT; 105 AA.
 AC P22574; 042369;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-B4 (ZF-13) (Fragment).
 GN HXB4 OR HXB4A OR HXB4B OR HXB4C OR ZF-13.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;

[1]
 SEQUENCE OF 1-86 FROM N.A.
 RP MEDLINE=89211958; PubMed=2468579;
 RA Noelstad P.R., Molven A., Eiken H.G., Fjose A.;
 RT "Structure and neural expression of a zebrafish homeobox sequence.";
 RL Gene 73:33-46(1988).
 [2]

SEQUENCE OF 44-105 FROM N.A.
 RP Prince V.E., Moens C.B., Kimmel C.B., Ho R.K.;
 RA "Zebrafish hox genes: expression in the hindbrain region of wild-type
 and mutants of the segmentation gene, valentino";
 RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 subfamily.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M24085; AAA56866.1; ALT_INIT.
 EMBL; Y13946; CAA74284.1; -.
 PIR; J70489; J70489.
 HSP; P02833; 9ANT.

ZFIN; ZDB-GENE-990415-105; hoxb4a.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 Transcription regulation.
 FT NON_TER 1 1
 FT DNA_BIND 10 69 HOMEBOX.
 SQ SEQUENCE 105 AA; 12262 MW; B0EFD84D909289F1 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRKWK 16
 DB 52 RQIKWIFQNRMRKWK 67

RESULT 19

HXB7 RAT STANDARD; PRT; 112 AA.
 AC P18864;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-B7 (R1B) (Fragment).
 GN HXB7 OR HOXB-7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;

[1]

SEQUENCE FROM N.A.
 RP STRAIN-Sprague-Dawley;
 RX MEDLINE=89231502; PubMed=2907739;
 RA Falzon M., Chung S.Y.;

RT "The expression of rat homeobox-containing genes is developmentally
 regulated and tissue specific.";
 RL Development 103:601-610(1988).

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M37566; AAA41342.2; -.
 HSP; P02833; 9ANT.

DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS50071; HOMEBOX 2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 Transcription regulation.

FT NON_TER 1 1

FT DNA_BIND 10 69 HOMEBOX.

SQ SEQUENCE 112 AA; 13910 MW; 099BGF064DC47C28 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 6.3e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 52 RQIKWIFQNRMKWK 67

RESULT 20

HXA5_AMEVE STANDARD; PRT; 148 AA.

ID HXA5_AMEVE STANDARD; PRT; 148 AA.

AC P50208;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Homeobox protein Hox-A5 (Fragment).

GN HOXA5.

OS Ambystoma mexicanum (Axolotl).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;

OC Ambystoma.

OC NCBI_TaxID=8296;

QY [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96032352; PubMed=7557438;

RA Gaur A.F., Lemanski L.F., Dube D.K.;

RT "Identification and expression of a homologue of the murine HoxA5

RT gene in the Mexican axolotl (Ambystoma mexicanum).";

RL Gene 162:249-253(1995).

CC -!- FUNCTION: Sequence-specific transcription factor which is part of

CC a developmental regulatory system that provides cells with

CC specific positional identities on the anterior-posterior axis (By

CC similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U19238; AAA91634.1; -.

DR PIR: PC4071; PC4071.

DR HSSP: P02833; ISAN.

DR TRANSFAC: T03305; -.

DR InterPro: IPR001827; Antennapedia.

DR InterPro: IPR001356; Homeobox.

DR Pfam: PF00046; homeobox; 1.

DR PRINTS: PR00025; ANTENNAPEDIA.

DR PRINTS: PR00024; HOMEBOX.

DR ProDom: PD000010; Homeobox; 1.

DR SMART: SM00389; HOX; 1.

DR PROSITE: PS00027; HOMEBOX 1; 1.

DR PROSITE: PS00032; ANTENNAPEDIA; 1.

DR PROSITE: PS00071; HOMEBOX 2; 1.

DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;

DR Transcription regulation.

DR NON_TER 1 1

FT DNA_BIND 73 132 HOMEBOX.

SQ SEQUENCE 148 AA; 16758 MW; C1893F0ED9BF5086 CRC64;

Query Match

Best Local Similarity 100.0%; Score 92; DB 1; Length 148;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16

|||||

Db 115 RQIKWIFQNRMKWK 130

RESULT 21

HXC6_SHEEP

ID HXC6_SHEEP STANDARD; PRT; 153 AA.

AC P49925;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-C6.
 GN HXC6.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 QY [1]
 RP SEQUENCE FROM N.A.
 RA Cosby N.C., Hernandez-Ledezma J., Mathialagan N., Roberts R.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U33049; AAA75473.1; -.
 DR HSSP: P02833; 9ANT.
 DR InterPro: IPR001827; Antennapedia.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00025; ANTENNAPEDIA.
 DR PRINTS: PR00024; HOMEBOX.
 DR ProDom: PD000010; Homeobox; 1.
 DR SMART: SM00389; HOX; 1.
 DR PROSITE: PS00027; HOMEBOX 1; 1.
 DR PROSITE: PS00032; ANTENNAPEDIA; 1.
 DR PROSITE: PS00071; HOMEBOX 2; 1.
 DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 DR Transcription regulation.
 DR SITE 40 45 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 59 118 HOMEBOX.
 FT DOMAIN 86 89 POLY-ARG.
 FT DOMAIN 126 130 POLY-GLY.
 FT DOMAIN 139 153 GLU/LYS-RICH.
 SQ SEQUENCE 153 AA; 17804 MW; 291E24399159621E CRC64;

Query Match 100.0%; Score 92; DB 1; Length 153;
 Best Local Similarity 100.0%; Pred. No. 8.7e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 |||||
 Db 101 RQIKWIFQNRMKWK 116

RESULT 22
 HXA7_HETFR
 ID HXA7_HETFR STANDARD; PRT; 208 AA.
 AC Q91A25;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-A7.
 GN HOXA7.
 OS Heterodontus francisci (Horn shark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
 OC Heterodontidae; Heterodontus.
 OC NCBI_TaxID=7792;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20144036; PubMed=10677514;
RA Kim C.B., Anemiyu C., Bailey W., Kawasaki K., Mezey J., Miller W.,
RT Minoshima S., Shimizu N., Wagner G., Ruddle F.;
RA "Hox cluster genomics in the horn shark, Heterodontus francisci.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).
CC -1- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP224262; AAF44645.1; -.
CC HSSP; P02833; 9ANT.
CC TRANSFAC; T04478; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS50071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC SITE 124 129 ANTP-TYPE HEXAPEPTIDE.
CC FT DNA BIND 135 194 HOMEBOX.
CC SQ SEQUENCE 208 AA; 23786 MW; 71B661AB78E064D0 CRC64;

Query Match 100.0%; Score 92; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16
Db 177 RQIKWIFQNRRMKWK 192

RESULT 23
HXA7 XENLA STANDARD; PRT; 209 AA.
AC P09071;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A7 (XlHox-3) (Xhox-36).
GN HXA7 OR XLHox3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8211489; PubMed=2452727;
RA Condie B.G., Harland R.M.;
RT "Posterior expression of a homeobox gene in early Xenopus embryos.";
RL Development 101:93-105(1987).
CC [2]
CC SEQUENCE OF 119-209 FROM N.A.
CC MEDLINE=88157707; PubMed=2894634;

```

```

RA Fritz A., de Robertis E.M.;
RT "Xenopus homeobox-containing cDNAs expressed in early development.";
RL Nucleic Acids Res. 16:1453-1469(1988).
CC -1- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED EXCLUSIVELY IN THE POSTERIOR
CC MESODERM AND ECTODERM OF EARLY XENOPUS EMBRYOS.
CC -1- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24752; AAA49753.1; -.
CC EMBL; X07103; CAA30124.1; AUT INIT.
CC PIR; A43553; A43553.
CC HSSP; P02833; 9ANT.
CC TRANSFAC; T01704; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS50071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC SITE 111 116 ANTP-TYPE HEXAPEPTIDE.
CC FT DNA BIND 122 181 HOMEBOX.
CC SQ SEQUENCE 209 AA; 23984 MW; 4EAC0A052F05D70D CRC64;

Query Match 100.0%; Score 92; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRRMKWK 16
Db 164 RQIKWIFQNRRMKWK 179

RESULT 24
HXB7 BOVIN STANDARD; PRT; 217 AA.
AC Q9TT89;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B7.
GN HXB7.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20302721; PubMed=10842316;
RA Bostrom K., Tintut Y., Kao S.C., Stanford W.P., Demer L.R.;
RT "HOXB7 overexpression promotes differentiation of C3H10T1/2 cells to
RT smooth muscle cells.";
RL J. Cell. Biochem. 78:210-221(2000).
CC -1- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:14:23 ; Search time 27.6757 Seconds

(without alignments)
182.409 Million cell updates/sec

Title: US-09-977-349-1

Perfect score: 92

Sequence: 1 RQIKWIFQNRMKWKX 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SPTREML_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	33	5 Q86FU0	Q86FU0 drosophila
2	92	100.0	39	13 O57368	O57368 brachydanio
3	92	100.0	42	11 Q80WH6	Q80WH6 rattus sp.
4	92	100.0	43	13 O57359	O57359 brachydanio
5	92	100.0	46	13 Q9PVR9	Q9PVR9 oryzias lat
6	92	100.0	51	5 Q27413	Q27413 ctenodrilus
7	92	100.0	51	5 Q23743	Q23743 ctenodrilus
8	92	100.0	51	5 Q26407	Q26407 ctenodrilus
9	92	100.0	57	13 Q9PVR8	Q9PVR8 oryzias lat
10	92	100.0	58	5 Q9Y188	Q9Y188 priapulus c
11	92	100.0	58	5 Q25208	Q25208 junonia coe
12	92	100.0	58	13 O57362	O57362 brachydanio
13	92	100.0	59	5 Q8WRM9	Q8WRM9 lithobius a
14	92	100.0	59	5 Q9NB42	Q9NB42 anopheles g
15	92	100.0	59	13 Q9PVR5	Q9PVR5 oryzias lat
16	92	100.0	60	5 O77143	O77143 archegozete

17	92	100.0	60	5 O77139	O77139 archegozete
18	92	100.0	60	11 Q80WH7	Q80WH7 rattus sp.
19	92	100.0	60	11 Q80WH4	Q80WH4 rattus sp.
20	92	100.0	60	13 Q8QGL5	Q8QGL5 petromyzon
21	92	100.0	60	13 Q8QGL3	Q8QGL3 petromyzon
22	92	100.0	60	13 Q8QGL6	Q8QGL6 petromyzon
23	92	100.0	60	13 Q8QGL2	Q8QGL2 petromyzon
24	92	100.0	60	13 Q8QGL8	Q8QGL8 petromyzon
25	92	100.0	60	13 Q8QGL7	Q8QGL7 petromyzon
26	92	100.0	61	5 Q27910	Q27910 polyandroca
27	92	100.0	61	5 O77138	O77138 archegozete
28	92	100.0	63	5 Q8MXB2	Q8MXB2 holopneuste
29	92	100.0	66	13 O57356	O57356 brachydanio
30	92	100.0	69	5 Q9U9T4	Q9U9T4 nereis vire
31	92	100.0	69	5 Q9BNF7	Q9BNF7 halloctis as
32	92	100.0	70	5 Q967W5	Q967W5 folsomia ca
33	92	100.0	70	13 Q801B4	Q801B4 latimeria m
34	92	100.0	71	13 Q9PVS3	Q9PVS3 oryzias lat
35	92	100.0	71	13 Q9PVS1	Q9PVS1 oryzias lat
36	92	100.0	73	5 Q9Y186	Q9Y186 priapulus c
37	92	100.0	73	5 Q86D93	Q86D93 spadeilla ce
38	92	100.0	74	13 O57367	O57367 brachydanio
39	92	100.0	75	5 Q25209	Q25209 junonia coe
40	92	100.0	75	13 Q9PVR6	Q9PVR6 oryzias lat
41	92	100.0	76	5 Q44257	Q44257 ethmostigm
42	92	100.0	76	5 Q86NB1	Q86NB1 calanus hel
43	92	100.0	77	5 Q44260	Q44260 ethmostigm
44	92	100.0	77	5 Q9Y187	Q9Y187 priapulus c
45	92	100.0	77	5 Q9U9Z4	Q9U9Z4 lingula ung
46	92	100.0	78	5 Q7Z0F3	Q7Z0F3 symagittif
47	92	100.0	79	5 Q967V2	Q967V2 lithobius f
48	92	100.0	79	5 Q9U9T9	Q9U9T9 nereis vire
49	92	100.0	80	5 Q05008	Q05008 artemia san
50	92	100.0	81	5 Q9BNZ7	Q9BNZ7 porcellio s
51	92	100.0	81	5 P91769	P91769 manduca sex
52	92	100.0	81	5 Q17142	Q17142 brachioisto
53	92	100.0	81	13 Q9PVR7	Q9PVR7 oryzias lat
54	92	100.0	82	5 Q24758	Q24758 drosophila
55	92	100.0	82	13 Q9PVS0	Q9PVS0 oryzias lat
56	92	100.0	82	13 Q801B8	Q801B8 latimeria m
57	92	100.0	83	5 Q17140	Q17140 brachioisto
58	92	100.0	85	13 Q801B9	Q801B9 latimeria m
59	92	100.0	86	13 Q9W7P8	Q9W7P8 paralichthy
60	92	100.0	87	13 Q8QFV4	Q8QFV4 lampetra fl
61	92	100.0	88	5 Q967W7	Q967W7 folsomia ca
62	92	100.0	89	5 Q967V4	Q967V4 lithobius f
63	92	100.0	90	5 Q94576	Q94576 helioctidari
64	92	100.0	90	5 Q817C9	Q817C9 ciona intes
65	92	100.0	92	13 O57377	O57377 brachydanio
66	92	100.0	93	5 Q9U9Z5	Q9U9Z5 lingula ung
67	92	100.0	93	5 Q8WQR9	Q8WQR9 eupyryna sc
68	92	100.0	94	5 Q967V6	Q967V6 lithobius f
69	92	100.0	96	5 Q9U9T8	Q9U9T8 nereis vire
70	92	100.0	96	5 Q62550	Q62550 lineus sang
71	92	100.0	97	5 Q9U9Z3	Q9U9Z3 lingula ung
72	92	100.0	98	5 Q8MX73	Q8MX73 sacculina c
73	92	100.0	100	5 Q9GYN5	Q9GYN5 tribolium c
74	92	100.0	100	5 Q8WQS0	Q8WQS0 eupyryna sc
75	92	100.0	100	5 Q8WQS1	Q8WQS1 eupyryna sc
76	92	100.0	101	13 Q9PDS5	Q9PDS5 gallus gall
77	92	100.0	103	5 Q62551	Q62551 lineus sang
78	92	100.0	105	5 Q17139	Q17139 brachioisto
79	92	100.0	106	5 Q26498	Q26498 schistocerc
80	92	100.0	107	5 Q8WRM8	Q8WRM8 lithobius a
81	92	100.0	108	11 Q8UJW2	Q8UJW2 mus musculu
82	92	100.0	109	5 Q94575	Q94575 helioctidari
83	92	100.0	109	5 Q05011	Q05011 artemia san
84	92	100.0	115	5 Q44258	Q44258 ethmostigm
85	92	100.0	115	13 Q801B5	Q801B5 latimeria m
86	92	100.0	120	5 Q44268	Q44268 acanthokara
87	92	100.0	140	13 Q801D0	Q801D0 latimeria m
88	92	100.0	146	5 Q05009	Q05009 artemia san
89	92	100.0	150	11 Q61679	Q61679 mus musculu

90 92 100.0 172 5 076844
 91 92 100.0 203 5 024681
 92 92 100.0 217 4 096BQ6
 93 92 100.0 218 5 09XY03
 94 92 100.0 219 13 090VZ9
 95 92 100.0 220 5 018313
 96 92 100.0 222 11 08BJW4
 97 92 100.0 223 13 0801A6
 98 92 100.0 224 13 08JH54
 99 92 100.0 225 13 0801C0
 100 92 100.0 227 5 017141

ALIGNMENTS

RESULT 1
 Q86FU0 PRELIMINARY; PRT; 33 AA.
 ID Q86FU0
 AC Q86FU0; (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Antennapedia complex (Fragment).
 GN ANT-C.
 OS Drosophila pseudoobscura (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7237;
 RN [1]_TaxID=7237;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93292933; PubMed=8099892;
 RA Randazzo F.M., Seeger M.A., Hues C.A., Sweeney M.A., Cecil J.K.,
 RA Kaufman T.C.;
 RT "Structural changes in the antennapedia complex of Drosophila
 RT pseudoobscura";
 RL Genetics 134:319-330(1993).
 DR EMBL; S63455; AAP13946.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3963 MW; D78E37ED81FD45DF CRC64;
 Query Match 100.0%; Score 92; DB 5; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKWQNRMKWK 16
 DB 7 RQIKWQNRMKWK 22
 RESULT 2
 O57368 PRELIMINARY; PRT; 39 AA.
 ID O57368
 AC O57368; (TREMELrel. 06, Created)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Hoxc5 protein (Fragment).
 GN HOXC5 OR HOXC5.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]_TaxID=7955;

RN SEQUENCE FROM N.A.
 RP Prince V.E., Joly L., Ekker M., Ho R.K.;
 RA "Zebrafish hox genes: genomic organization and modified colinear
 RT expression patterns in the trunk";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; Y14539; CAA74874.1; -.
 DR ZFIN; ZDB-GENE-980526-533; hoxc5a.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambrepresr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 39 39
 SQ SEQUENCE 39 AA; 4827 MW; 592A0FEC12B58860 CRC64;
 Query Match 100.0%; Score 92; DB 13; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RQIKWQNRMKWK 16
 DB 14 RQIKWQNRMKWK 29
 RESULT 3
 Q80WH6 PRELIMINARY; PRT; 42 AA.
 ID Q80WH6
 AC Q80WH6;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hox-A Hox-1 (Fragment).
 GN HOX-A HOX-1.
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10118;
 RN [1]_TaxID=10118;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95217128; PubMed=7702549;
 RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
 RT "Cloning of rat homeobox genes";
 RL Biochem. Genet. 32:351-360(1994).
 DR EMBL; S76290; AAP31864.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambrepresr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 FT NON_TER 1 1
 FT NON_TER 42 42
 SQ SEQUENCE 42 AA; 5494 MW; 38E5153B92216FE9 CRC64;
 Query Match 100.0%; Score 92; DB 11; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 DB 25 RQIKWIFQNRMKWK 40
 RESULT 4
 OS7359 PRELIMINARY; PRT; 43 AA.
 AC O57359;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hoxa5 protein (Fragment).
 GN HOXB5 OR HOXA5
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Prince V.E., Joly L., Ekker M., Ho R.K.;
 RT "Zebrafish hox genes: genomic organization and modified colinear
 RT expression patterns in the trunk."
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; Y14526; CAA74861.1; -;
 DR ZFIN; ZDB-GENE-000823-6; hox5b.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 43 AA; 5050 MW; 53034C37F3DFA596 CRC64;

Query Match 100.0%; Score 92; DB 13; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 DB 10 RQIKWIFQNRMKWK 25

RESULT 5
 Q9PVR9 PRELIMINARY; PRT; 46 AA.
 AC Q9PVR9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE HOXC5A (Fragment).
 GN HOXC5A.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kondo S., Naruse K., Shima A.;
 RT "Hox genes of the medakafish Oryzias latipes."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AB026960; BAA86243.1; -;

HSP; P02833; IHOM.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 46 AA; 5955 MW; 6039999ED4294DD3 CRC64;

Query Match 100.0%; Score 92; DB 13; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMKWK 16
 DB 23 RQIKWIFQNRMKWK 38

RESULT 6
 Q27413 PRELIMINARY; PRT; 51 AA.
 AC Q27413;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE LOK5 ORTHOLOG homeobox (Fragment).
 GN CTS-LOX5.
 OS Ctenodrilus serratus.
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
 OC Terebellida; Ctenodrilidae; Ctenodrilus.
 OX NCBI_TaxID=40316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94356262; PubMed=7915607;
 RA Dick M.H., Buss L.W.;
 RT "A PCR-based survey of homeobox genes in Ctenodrilus serratus
 RT (Annelida: Polychaeta)."
 RL Mol. Phylogenet. Evol. 3:146-158 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dick M.H., Buss L.W.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; U26629; AAC46851.1; -;
 DR EMBL; S76226; AAB31777.1; -;
 DR HSP; P02833; 9ANT.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 51 AA; 6278 MW; 88C8F65161E94A22 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 51;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 ROIKWQFQNRKMKKK 16
DB      23 ROIKWQFQNRKMKKK 38

RESULT 7
Q23743
ID      Q23743      PRELIMINARY;      PRT;      51 AA.
AC      Q23743      (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Deformed ortholog homeobox (Fragment).
GN      CTS-DFD.
OS      Ctenodrilus serratus.
OC      Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC      Terebellida; Ctenodrilidae; Ctenodrilus.
OX      NCBI_TaxID=40316;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94356262; PubMed=7915607;
RA      Dick M.H., Buss L.W.;
RT      "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RL      (Annelida: Polychaeta).";
RN      Mol. Phylogenet. Evol. 3:146-158(1994).
RL      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94356262; PubMed=7915607;
RA      Dick M.H., Buss L.W.;
RT      Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC      -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR      EMBL; U26627; AAC46849.1; -.
DR      HSSP; P02833; 9ANT.
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      InterPro; IPR001356; Homeobox.
DR      Pfam; PR00046; homeobox; 1.
DR      PRINTS; PR00024; HOMEBOX.
DR      PRINTS; PR00031; HTHREPRESSR.
DR      ProDom; PD000010; Homeobox; 1.
DR      PROSITE; PS00027; HOMEBOX_1; 1.
DR      PROSITE; PS50071; HOMEBOX_2; 1.
KW      DNA-binding; Homeobox; Nuclear protein.
FT      NON_TER      1
FT      NON_TER      51
SQ      SEQUENCE      51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match      100.0%; Score 92; DB 5; Length 51;
Best Local Similarity      100.0%; Pred. No. 2e-07;
Matches      16; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 ROIKWQFQNRKMKKK 16
DB      23 ROIKWQFQNRKMKKK 38

RESULT 8
Q26407
ID      Q26407      PRELIMINARY;      PRT;      51 AA.
AC      Q26407;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Cts-Dfd protein (Fragment).
GN      CTS-DFD.
OS      Ctenodrilus serratus.
OC      Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC      Terebellida; Ctenodrilidae; Ctenodrilus.
OX      NCBI_TaxID=40316;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94356262; PubMed=7915607;

```

```

RA      Dick M.H., Buss L.W.;
RT      "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RL      (Annelida: Polychaeta).";
RN      Mol. Phylogenet. Evol. 3:146-158(1994).
CC      -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR      EMBL; S76416; AAB31775.1; -.
DR      HSSP; P02833; 9ANT.
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      InterPro; IPR001356; Homeobox.
DR      InterPro; IPR000047; HTH lambrepressr.
DR      Pfam; PR00046; homeobox; 1.
DR      PRINTS; PR00024; HOMEBOX.
DR      PRINTS; PR00031; HTHREPRESSR.
DR      ProDom; PD000010; Homeobox; 1.
DR      SMART; SM00389; HOX; 1.
DR      PROSITE; PS00027; HOMEBOX_1; 1.
DR      PROSITE; PS50071; HOMEBOX_2; 1.
KW      DNA-binding; Homeobox; Nuclear protein.
FT      NON_TER      1
FT      NON_TER      51
SQ      SEQUENCE      51 AA; 6533 MW; 9EDB50C927FBCBD5 CRC64;

Query Match      100.0%; Score 92; DB 5; Length 51;
Best Local Similarity      100.0%; Pred. No. 2e-07;
Matches      16; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 ROIKWQFQNRKMKKK 16
DB      23 ROIKWQFQNRKMKKK 38

RESULT 9
Q26407
ID      Q26407      PRELIMINARY;      PRT;      57 AA.
AC      Q26407;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      HOXA5A (Fragment).
GN      HOXA5A.
OS      Oryzias latipes (Medaka fish) (Japanese ricefish).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC      Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX      NCBI_TaxID=8090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      Kondo S., Naruse K., Shima A.;
RT      "Hox genes of the medakafish Oryzias latipes.";
RL      Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC      -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR      EMBL; AB026961; EAA86244.1; -.
DR      HSSP; P02833; IHOM.
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0003700; F:transcription factor activity; IEA.
DR      GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR      InterPro; IPR001356; Homeobox.
DR      Pfam; PR00046; homeobox; 1.
DR      PRINTS; PR00024; HOMEBOX.
DR      ProDom; PD000010; Homeobox; 1.
DR      SMART; SM00389; HOX; 1.
DR      PROSITE; PS00027; HOMEBOX_1; 1.
DR      PROSITE; PS50071; HOMEBOX_2; 1.
KW      DNA-binding; Homeobox; Nuclear protein.
FT      NON_TER      1
FT      NON_TER      57
SQ      SEQUENCE      57 AA; 6891 MW; 54A6430320F68C04 CRC64;

Query Match      100.0%; Score 92; DB 13; Length 57;
Best Local Similarity      100.0%; Pred. No. 2.3e-07;
Matches      16; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY 1 RQIKWIFQNRRMKWK 16
DB 23 RQIKWIFQNRRMKWK 38

RESULT 10
QY188 PRELIMINARY; PRT; 58 AA.
AC Q9Y188
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HBI homeodomain protein (Fragment).
GN HBI.
OS Priapulius caudatus.
OC Eukaryota; Metazoa; Priapulida; Priapulidae; Priapulidae; Priapulidae.
OX NCBI_TaxID=37621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99318125; PubMed=10391241;
RA de Rosa R., Grenier J.K., Andreeva T., Cook C.E., Adoutte A., Akam M.,
RA Carroll S.B., Balavoine G.;
RT "Hox genes in brachiopods and priapulids and protostome evolution.";
RL Nature 399:772-776(1999).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF144888; AAD40644.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1 1
FT SEQUENCE 58 AA; 7323 MW; 572F30DA57C9A613 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
DB 24 RQIKWIFQNRRMKWK 39

RESULT 11
Q25208 PRELIMINARY; PRT; 58 AA.
AC Q25208;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Antennapedia protein (Fragment).
GN ANTENNAPEIDIA.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95075456; PubMed=7840822;
RA Warren R.W., Nagy L., Selegue J., Gates J., Carroll S.;
RA "Evolution of homeotic gene regulation and function in flies and
RT butterflies.";
RL Nature 372:458-461(1994).

```

```

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; L42135; AAA68461.1; -.
DR PIR; S58850; S58850.
DR HSSP; P02833; 1HOM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambrepreser.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1 1
FT SEQUENCE 58 AA; 7583 MW; BD69B4875BAE565E CRC64;

Query Match 100.0%; Score 92; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
DB 43 RQIKWIFQNRRMKWK 58

RESULT 12
O57362 PRELIMINARY; PRT; 58 AA.
AC O57362;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hoxb7 protein (Fragment).
GN HOXB7A OR HOXB7.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Prince V.E., Joly L., Ekker M., Ho R.K.;
RT "Zebrafish hox genes: genomic organization and modified colinear
RT expression patterns in the trunk.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; Y14533; CAA74868.1; -.
DR TRANSFAC; T01635; -.
DR ZFIN; ZDB-GENE-000129-2; hoxb7a.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambrepreser.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1 1
FT SEQUENCE 58 AA; 6863 MW; 41EA6F14488DA1E5 CRC64;

Query Match 100.0%; Score 92; DB 13; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 RQIKWFFQNRMMKWK 16
Db 14 RQIKWFFQNRMMKWK 29

RESULT 13
Q9WRM9
ID Q9WRM9 PRELIMINARY; PRT; 59 AA.
AC Q9WRM9;
DT 01-WAR-2002 (TREMELrel. 20, Created)
DT 01-WAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Antennapedia (Fragment).
GN ANTP.
OS Lithobius atkinsoni.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=177213;
RN [1]
RP SEQUENCE FROM N.A.
RA Hughes C.L., Kaufman T.C.;
RT "Exploring the myriapod body plan: expression patterns of the ten Hox
genes in a centipede";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF434996; AAL36901.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 59 AA; 7093 MW; 9B60036CE0D515C1 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFFQNRMMKWK 16
Db 23 RQIKWFFQNRMMKWK 38

RESULT 14
Q9NB42
ID Q9NB42 PRELIMINARY; PRT; 59 AA.
AC Q9NB42;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Transcription factor deformed (Fragment).
GN DFD.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21150895; PubMed=11256376;
RA Powers T.P., Hogan J., Ke Z., Dymbrowski K., Wang X., Collins F.H.,
RA Kaufman T.C.;
RT "Characterization of the Hox cluster from the mosquito Anopheles
gambiae (Diptera: Culicidae).";
RL Evol. Dev. 2:311-325(2000).

```

```

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF269155; AAF91400.1; -.
DR HSSP; P02833; ISAN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 59 AA; 7621 MW; C38A2505A81D9952 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFFQNRMMKWK 16
Db 43 RQIKWFFQNRMMKWK 58

RESULT 15
Q9PVR5
ID Q9PVR5 PRELIMINARY; PRT; 59 AA.
AC Q9PVR5;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE HOXB6B (Fragment).
GN HOXB6B.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RT "Hox genes of the medakafish Oryzias latipes.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB026964; BAA86247.1; -.
DR HSSP; P02833; IHOM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS0071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 59 AA; 7425 MW; AA2181EC56B968D CRC64;

Query Match 100.0%; Score 92; DB 13; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFFQNRMMKWK 16
Db 23 RQIKWFFQNRMMKWK 38

```

```

RESULT 16
O77143 ID O77143 PRELIMINARY; PRT; 60 AA.
AC O77143;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Sex combs reduced (Fragment).
GN SCR.
OS Archegozetes longisetosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcopiformes; Oribatida; Desmonomata;
OC Trhypochthonioidae; Trhypochthoniidae; Archegozetes.
OX NCBI_taxID=66560;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98393703; PubMed=9724762;
RA Telford M.J., Thomas R.H.;
RT "Expression of homeobox genes shows chelicerate arthropods retain
their deutocerebral segment.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10671-10675(1998).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR HSSP; P02833; 1SAN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000047; HTH_lambrepresr.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7755 MW; 32678A250BDEF74 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRKWK 16
Db 43 RQIKWIFQNRMRKWK 58

RESULT 17
O77139 ID O77139 PRELIMINARY; PRT; 60 AA.
AC O77139;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Deformed (Fragment).
GN DFD.
OS Archegozetes longisetosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcopiformes; Oribatida; Desmonomata;
OC Trhypochthonioidae; Trhypochthoniidae; Archegozetes.
OX NCBI_taxID=66560;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98393703; PubMed=9724762;
RA Telford M.J., Thomas R.H.;
RT "Expression of homeobox genes shows chelicerate arthropods retain
their deutocerebral segment.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10671-10675(1998).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR HSSP; AF071403; AAC35932.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7807 MW; 0A22E87CA4C98143 CRC64;

Query Match 100.0%; Score 92; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRKWK 16
Db 43 RQIKWIFQNRMRKWK 58

RESULT 18
O80WH7 ID O80WH7 PRELIMINARY; PRT; 60 AA.
AC O80WH7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hox-A Hox-1 (Fragment).
GN HOX-A HOX-1.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxID=10118;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RL Biochem. Genet. 32:351-360(1994).
DR EMBL; S76287; AAP31863.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7849 MW; 1334ED69B4875BAE CRC64;

Query Match 100.0%; Score 92; DB 11; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWIFQNRMRKWK 16
Db 43 RQIKWIFQNRMRKWK 58

RESULT 19

```

```

Q80WH4 ID Q80WH4 PRELIMINARY; PRT; 60 AA.
AC Q80WH4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hox-B|Hox-2 (Fragment)
GN Hox-B|Hox-2.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes";
RL Biochem. Genet. 32:351-360(1994).
DR EMBL; S76297; AAP31867.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7842 MW; 3407ED69B4874816 CRC64;

Query Match 100.0%; Score 92; DB 11; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 43 RQIKWIFQNRMRMKWK 58

RESULT 20
Q80GL5 ID Q80GL5 PRELIMINARY; PRT; 60 AA.
AC Q80GL5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HoxL5/6 homeobox (Fragment)
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX Irvine S.O., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
RA Amemiya C.T., Ruddle F.H.;
RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
marinus.";
RL J. Exp. Zool. 0:0-0(2002).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF410914; AAM19472.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7863 MW; 1334F475B49B47FE CRC64;

Query Match 100.0%; Score 92; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 43 RQIKWIFQNRMRMKWK 58

RESULT 22
Q80GL6 ID Q80GL6 PRELIMINARY; PRT; 60 AA.
AC Q80GL6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HoxK6 homeobox (Fragment)

```

```

DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7717 MW; E150F172FD751F3C CRC64;

Query Match 100.0%; Score 92; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 43 RQIKWIFQNRMRMKWK 58

RESULT 21
Q80GL3 ID Q80GL3 PRELIMINARY; PRT; 60 AA.
AC Q80GL3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HoxN6 homeobox (Fragment)
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX Irvine S.O., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
RA Amemiya C.T., Ruddle F.H.;
RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
marinus.";
RL J. Exp. Zool. 0:0-0(2002).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF410916; AAM19474.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7863 MW; 1334F475B49B47FE CRC64;

Query Match 100.0%; Score 92; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 43 RQIKWIFQNRMRMKWK 58

RESULT 22
Q80GL6 ID Q80GL6 PRELIMINARY; PRT; 60 AA.
AC Q80GL6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HoxK6 homeobox (Fragment)

```

FT	NON_TER	60
SQ	SEQUENCE	60 AA; 7863 MW; 1334F475E49E47FE CRC64;
 Query Match 100.0%; Score 92; DB 13; Length 60; Best Local Similarity 100.0%; Pred. No. 2.4e-07; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 RQIKWIFQNRRMKWK 16 	
Db	43 RQIKWIFQNRRMKWK 58 	
 RESULT 24 SQ8QLG8 PRELIMINARY; PRT; 60 AA.		
ID	AC Q8QGL8;	
AD	DT 01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DE	DE HoxG4 homeobox (Fragment).	
OS	OS Petromyzon marinus (Sea lamprey).	
OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;	
OC	OC Petromyzontiformes; Petromyzontidae; Petromyzon.	
OX	NCBI_TaxID=7757;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Irvine S.O., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,	
RA	Ameniya C.T., Ruddle F.H.;	
RT	"Genomic analysis of Hox clusters in the sea lamprey Petromyzon	
RT	marinus";	
RL	J. Exp. Zool. 0:0-0(2002).	
RC	- - SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).	
DR	EMB1; AF410911; AAM19469.1; -	
DR	GO; GO:0005634; C:nucleus; IEA.	
DR	GO; GO:0003700; F:transcription factor activity; IEA.	
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001356; Homeobox.	
DR	InterPro; IPR000047; HTH_lambdarepressr.	
DR	Pfam; PF00046; homeobox_1.	
DR	PRINTS; PRK0024; HOMEBOX.	
DR	PRINTS; PRK0031; HTREPRESSR.	
DR	ProDom; PD000010; Homeobox; 1.	
DR	SMART; SM00389; HOX; 1.	
DR	PROSITE; PS00027; HOMEBOX_1; 1.	
DR	PROSITE; PS00071; HOMEBOX_2; 1.	
KW	Homeobox; DNA-binding; Nuclear protein.	
FT	NON_TER	1
FT	NON_TER	60
SQ	SEQUENCE	60 AA; 7704 MW; 0EEB3642C24DEC6E CRC64;
 Query Match 100.0%; Score 92; DB 13; Length 60; Best Local Similarity 100.0%; Pred. No. 2.4e-07; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 RQIKWIFQNRRMKWK 16 	
Db	43 RQIKWIFQNRRMKWK 58 	
 RESULT 25 SQ8QLG7 PRELIMINARY; PRT; 60 AA.		
ID	AC Q8QGL7;	
AD	DT 01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DE	DE HoxJ5 homeobox (Fragment).	
OS	OS Petromyzon marinus (Sea lamprey).	
OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;	
OC	OC Petromyzontiformes; Petromyzontidae; Petromyzon.	
OX	NCBI_TaxID=7757;	
RN	[1]	
RP	SEQUENCE FROM N.A.	

RA Irvine S.O., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
RA Anemiyu C.T., Ruddle F.H.;
RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
RT marinus";
RL J. Exp. Zool. 0:0-0 (2002).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF410912; AM19470.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; ED000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1 60
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 7635 MW; E1F1209C5ACDF866 CRC64;

Query Match 100.0%; Score 92; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRKWK 16
|||
Db 43 RQIKWIFQNRMRKWK 58

Search completed: May 24, 2004, 17:24:38
JOB time : 31.6757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:22:27 ; Search time 11.4595 Seconds
(without alignments)
72.082 Million cell updates/sec

Title: US-09-977-349-1

Perfect score: 92

Sequence: 1 RQIKWFKQRNKKWK 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	16	2	US-08-928-958-7
2	92	100.0	16	2	US-08-810-540-3
3	92	100.0	16	2	US-08-810-540-6
4	92	100.0	16	2	US-09-072-429-7
5	92	100.0	16	3	US-08-964-302A-6
6	92	100.0	16	3	US-08-116-294-4
7	92	100.0	16	3	US-08-964-614A-4
8	92	100.0	16	3	US-08-849-486-1
9	92	100.0	16	3	US-08-849-486-4
10	92	100.0	16	3	US-09-208-966-54
11	92	100.0	16	3	US-09-308-935-8
12	92	100.0	16	3	US-08-441-416A-6
13	92	100.0	16	4	US-09-296-089-33
14	92	100.0	16	4	US-09-419-826-35
15	92	100.0	16	4	US-09-302-305C-10
16	92	100.0	16	4	US-09-346-847-1
17	92	100.0	16	4	US-09-346-847-25
18	92	100.0	16	4	US-09-057-363C-47
19	92	100.0	16	4	US-09-043-560B-3
20	92	100.0	16	4	US-09-648-400A-29
21	92	100.0	16	4	US-09-227-652B-4
22	92	100.0	16	4	US-09-780-070-38
23	92	100.0	16	4	US-08-610-220B-9
24	92	100.0	16	4	US-09-775-052A-54
25	92	100.0	16	4	US-09-155-165-22
26	92	100.0	16	4	US-09-792-480-29
27	92	100.0	16	4	US-09-551-976-33

SUMMARIES

28	92	100.0	16	4	US-09-265-107-47	Sequence 47, Appl
29	92	100.0	17	4	US-09-346-847-17	Sequence 17, Appl
30	92	100.0	17	4	US-09-346-847-20	Sequence 22, Appl
31	92	100.0	17	4	US-09-346-847-22	Sequence 20, Appl
32	92	100.0	17	4	US-09-346-847-27	Sequence 27, Appl
33	92	100.0	17	4	US-09-648-400A-30	Sequence 30, Appl
34	92	100.0	18	3	US-08-838-545-20	Sequence 20, Appl
35	92	100.0	18	3	US-09-349-532-20	Sequence 20, Appl
36	92	100.0	19	4	US-09-346-847-23	Sequence 23, Appl
37	92	100.0	19	4	US-09-658-517C-7	Sequence 7, Appl
38	92	100.0	20	4	US-09-346-847-16	Sequence 16, Appl
39	92	100.0	20	4	US-09-346-847-18	Sequence 18, Appl
40	92	100.0	20	4	US-09-346-847-30	Sequence 30, Appl
41	92	100.0	20	4	US-09-658-517C-8	Sequence 8, Appl
42	92	100.0	21	4	US-08-610-220B-11	Sequence 11, Appl
43	92	100.0	22	4	US-09-346-847-28	Sequence 28, Appl
44	92	100.0	22	4	US-09-057-363C-50	Sequence 50, Appl
45	92	100.0	22	4	US-08-610-220B-10	Sequence 10, Appl
46	92	100.0	22	4	US-09-155-165-5	Sequence 5, Appl
47	92	100.0	22	4	US-09-155-165-11	Sequence 11, Appl
48	92	100.0	22	4	US-09-265-107-50	Sequence 50, Appl
49	92	100.0	24	4	US-09-419-826-34	Sequence 34, Appl
50	92	100.0	24	4	US-09-428-082B-332	Sequence 332, App
51	92	100.0	27	3	US-09-051-934-51	Sequence 51, Appl
52	92	100.0	27	3	US-09-051-934-52	Sequence 52, Appl
53	92	100.0	27	4	US-09-040-725A-2	Sequence 2, Appl
54	92	100.0	34	4	US-09-347-504-79	Sequence 79, Appl
55	92	100.0	34	4	US-10-161-499-79	Sequence 79, Appl
56	92	100.0	36	4	US-09-428-082B-331	Sequence 331, App
57	92	100.0	61	2	US-08-202-044-3	Sequence 3, Appl
58	92	100.0	61	3	US-08-751-344B-3	Sequence 3, Appl
59	92	100.0	61	3	US-08-751-344B-6	Sequence 6, Appl
60	92	100.0	61	3	US-08-751-344B-9	Sequence 9, Appl
61	91	98.9	61	3	US-08-751-344B-7	Sequence 7, Appl
62	87	94.6	15	2	US-08-810-540-4	Sequence 4, Appl
63	87	94.6	16	4	US-09-792-480-30	Sequence 30, Appl
64	87	94.6	42	3	US-08-751-344B-4	Sequence 4, Appl
65	87	94.6	283	1	US-08-583-672-2	Sequence 2, Appl
66	87	94.6	283	2	US-08-202-044-2	Sequence 2, Appl
67	87	94.6	283	3	US-08-751-344B-2	Sequence 2, Appl
68	87	94.6	284	2	US-08-320-148B-2	Sequence 2, Appl
69	87	94.6	284	3	US-08-589-028-6	Sequence 6, Appl
70	87	94.6	284	3	US-08-784-582-6	Sequence 6, Appl
71	87	94.6	284	3	US-08-785-271-6	Sequence 6, Appl
72	87	94.6	284	3	US-09-031-898-2	Sequence 2, Appl
73	87	94.6	302	2	US-08-203-532F-4	Sequence 4, Appl
74	87	94.6	302	3	US-08-950-860-16	Sequence 16, Appl
75	87	94.6	302	3	US-09-078-465-4	Sequence 4, Appl
76	87	94.6	302	5	PCT-US95-01882A-4	Sequence 4, Appl
77	87	94.6	303	2	US-08-203-532F-2	Sequence 2, Appl
78	87	94.6	303	2	US-09-078-465-2	Sequence 2, Appl
79	87	94.6	303	5	PCT-US95-01882A-2	Sequence 2, Appl
80	86	93.5	16	3	US-08-849-486-5	Sequence 5, Appl
81	86	93.5	16	4	US-09-466-772-7	Sequence 7, Appl
82	86	93.5	20	4	US-09-466-772-3	Sequence 3, Appl
83	86	93.5	21	4	US-09-466-772-1	Sequence 1, Appl
84	86	93.5	22	4	US-09-466-772-2	Sequence 2, Appl
85	86	93.5	23	4	US-08-466-772-4	Sequence 4, Appl
86	83	90.2	16	3	US-08-849-486-8	Sequence 8, Appl
87	83	90.2	300	4	US-09-162-524-1	Sequence 1, Appl
88	83	90.2	349	4	US-09-162-524-3	Sequence 3, Appl
89	82	89.1	16	4	US-09-402-929-7	Sequence 7, Appl
90	80	87.0	16	3	US-08-849-486-7	Sequence 7, Appl
91	79	85.9	61	2	US-08-757-316C-30	Sequence 30, Appl
92	78	84.8	16	3	US-09-208-966-1	Sequence 1, Appl
93	78	84.8	16	4	US-09-775-052A-1	Sequence 1, Appl
94	78	84.8	342	4	US-09-165-827C-2	Sequence 2, Appl
95	77	83.7	60	3	US-08-751-344B-5	Sequence 5, Appl
96	76	82.6	367	4	US-09-009-816-2	Sequence 2, Appl
97	75	81.5	16	4	US-09-396-089-34	Sequence 34, Appl
98	75	81.5	16	4	US-09-336-093-3	Sequence 3, Appl
99	75	81.5	16	4	US-09-057-363C-48	Sequence 48, Appl
100	75	81.5	16	4	US-09-557-465D-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-08-928-958-7
; Sequence 7, Application US/08928958
; Patent No. 5877282
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CLEVELAND, JEFFREY S.
; APPLICANT: BLAKE, JAMES
; APPLICANT: HAFAR, OMAR K.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN
; TITLE OF INVENTION: TRANSLLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,958
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 60/026978
; FILING DATE: 20-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-928-958-7

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRKMKWK 16
Db 1 RQIKWIFQNRKMKWK 16

RESULT 2

US-08-810-540-3
; Sequence 3, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
; APPLICANT: TROY, CAROL M.
; APPLICANT: Shelanski, Michael L.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,540
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526

ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,540
FILING DATE: 03-MAR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/51247
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-810-540-3

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRKMKWK 16
Db 1 RQIKWIFQNRKMKWK 16

RESULT 3

US-08-810-540-6
; Sequence 6, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
; APPLICANT: TROY, CAROL M.
; APPLICANT: Shelanski, Michael L.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,540
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526

Query Match 100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRKMKWK 16
Db 1 RQIKWIFQNRKMKWK 16

```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-810-540-6

Query Match      100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 4
US-09-072-429-7
; Sequence 7, Application US/09072429
; Patent No. 5962415
; GENERAL INFORMATION:
; APPLICANT: Nadler, Steven G.
; TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
; TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
; TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072.429
; FILING DATE: 04-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Klein, Christopher A.
; REGISTRATION NUMBER: 34,363
; REFERENCE/DOCKET NUMBER: ON0141b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-3714
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-072-429-7

Query Match      100.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 5
US-08-964-302A-6
; Sequence 6, Application US/08964302A
; Patent No. 6015787
```

```
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,302A
; FILING DATE: 04-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 00398/126001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-302A-6

Query Match      100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16

RESULT 6
US-09-116-294-4
; Sequence 4, Application US/09116294
; Patent No. 6025140
; GENERAL INFORMATION:
; APPLICANT: Langel, Ulo
; APPLICANT: Bartfal, Tamas
; APPLICANT: Pooga, Margus
; APPLICANT: Valkna, Andres
; APPLICANT: Saar, Kulliki
; APPLICANT: Hallbrink, Mattias
; TITLE OF INVENTION: Conjugated Constructs of Peptides and
; TITLE OF INVENTION: Nucleic Acid Analogs, and their Transport Across Membranes
; FILE REFERENCE: 4394
; CURRENT APPLICATION NUMBER: US/09/116,294
; CURRENT FILING DATE: 1998-07-16
; EARLIER APPLICATION NUMBER: 60/052,678
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: drosophila
US-09-116-294-4

Query Match      100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
```

```
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16
RESULT 7
US-08-964-614A-4
; Sequence 4, Application US/08964614A
; Patent No. 6057104
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN
; TITLE OF INVENTION: Rad51 PROTEIN AND DISRUPTION OF PROTEINS THAT ASSOCIATE
; TITLE OF INVENTION: WITH MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,614A
; FILING DATE: 05-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,280
; FILING DATE: 05-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8535-0019-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-614A-4
Query Match 100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16
RESULT 8
US-08-849-486-1
; Sequence 1, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-849-486-1
Query Match 100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16
RESULT 9
US-08-849-486-4
; Sequence 4, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..16
; OTHER INFORMATION: /product= "amino acids of the D series"
US-08-849-486-4
Query Match 100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RQIKWQNRRMKWKK 16
Db 1 RQIKWQNRRMKWKK 16
RESULT 10
US-09-208-966-54
```

```

; Sequence 54, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 4881/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 54
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-208-966-54

```

```

Query Match      100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RQIKWIFQNRMRKWK 16
Db      1 RQIKWIFQNRMRKWK 16

```

```

RESULT 11
US-09-308-935-8
; Sequence 8, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-308-935-8

```

```

Query Match      100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RQIKWIFQNRMRKWK 16
Db      1 RQIKWIFQNRMRKWK 16

```

```

RESULT 12
US-09-441-416A-6
; Sequence 6, Application US/09441416A
; Patent No. 6294518
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF
; TITLE OF INVENTION: CALPAIN
; FILE REFERENCE: 00398-140001
; CURRENT APPLICATION NUMBER: US/09/441,416A
; CURRENT FILING DATE: 1999-11-16

```

```

; PRIOR APPLICATION NUMBER: US 08/964,302
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-441-416A-6

```

```

Query Match      100.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RQIKWIFQNRMRKWK 16
Db      1 RQIKWIFQNRMRKWK 16

```

```

RESULT 13
US-09-296-089-33
; Sequence 33, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
; FILE REFERENCE: 100086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-296-089-33

```

```

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 RQIKWIFQNRMRKWK 16
Db      1 RQIKWIFQNRMRKWK 16

```

```

RESULT 14
US-09-419-826-35
; Sequence 35, Application US/09419826
; Patent No. 6306832
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDE ANTITESTROGEN COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR TREATING BREAST CANCER
; NUMBER OF SEQUENCES: 39
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/419,826
; FILING DATE: 14-OCT-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/07711
; FILING DATE: 14-APR-1998
; APPLICATION NUMBER: US 60/043,545
; FILING DATE: 14-APR-1997
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-419-826-35

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 15
US-09-302-305C-10
; Sequence 10, Application US/09302305C
; Patent No. 6350572
; GENERAL INFORMATION:
; APPLICANT: Bernards, Rene
; APPLICANT: Zwijssen, Renate
; TITLE OF INVENTION: Interaction Between Cyclin D1 and Steroid Receptor
; FILE REFERENCE: 4238/80713
; CURRENT APPLICATION NUMBER: US/09/302,305C
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/GB99/00440
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Translocation peptide derived from antennapedia
; OTHER INFORMATION: homeodomain protein
US-09-302-305C-10

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 16
US-09-346-847-1
; Sequence 1, Application US/09346847
; Patent No. 6472507
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/09/346,847
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-346-847-1

Query Match      100.0%; Score 92; DB 4; Length 16;

```

```

Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 17
US-09-346-847-25
; Sequence 25, Application US/09346847
; Patent No. 6472507
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/09/346,847
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: MOD RES
; LOCATION: (16)
; OTHER INFORMATION: AMIDATION
US-09-346-847-25

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 18
US-09-057-363C-47
; Sequence 47, Application US/09057363C
; Patent No. 6551994
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING THE
; INTERACTION BETWEEN ALPHA-CATENIN AND BETA-CATENIN
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,363C
; FILING DATE: 08-Apr-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614

```

; REFERENCE/DOCKET NUMBER: 100086.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-057-362C-47

Query Match 100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 19
US-09-043-560B-3
; Sequence 3, Application US/09043560B
; Patent No. 6569833
; GENERAL INFORMATION:
; APPLICANT: Fabraeus, Robin
; APPLICANT: Lane, David P.
; TITLE OF INVENTION: Cyclin Dependent Kinase Binding Compounds
; FILE REFERENCE: CCI-003US
; CURRENT APPLICATION NUMBER: US/09/043.560B
; CURRENT FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-043-560B-3

Query Match 100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 20
US-09-648-400A-29
; Sequence 29, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; FILE REFERENCE: 019801-000210US
; CURRENT APPLICATION NUMBER: US/09/648,400A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 29
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Antennapedia
; OTHER INFORMATION: homeodomain, Antennapedia-43-58
US-09-648-400A-29

Query Match 100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 21
US-09-227-652B-4
; Sequence 4, Application US/09227652B
; Patent No. 6610495
; GENERAL INFORMATION:
; APPLICANT: TVW Telethon Institute for Child Health Research
; TITLE OF INVENTION: PEPTIDE DETECTION METHOD
; FILE REFERENCE: 1991209/MRO-PCT
; CURRENT APPLICATION NUMBER: US/09/227.652B
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 60/070989
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Penetratin 16-mer.
US-09-227-652B-4

Query Match 100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRRMKWK 16
Db 1 RQIKWIFQNRRMKWK 16

RESULT 22
US-09-780-070-38
; Sequence 38, Application US/09780070
; Patent No. 6632616
; GENERAL INFORMATION:
; APPLICANT: Burke, James
; APPLICANT: Strittmatter, Warren
; APPLICANT: Nagai, Yoshitaka
; TITLE OF INVENTION: COMPOUNDS THAT SELECTIVELY BIND TO EXPANDED POLYGLUTAMINE REPEAT
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 5405,242
; CURRENT APPLICATION NUMBER: US/09/780,070
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/189,781
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-780-070-38

Query Match 100.0%; Score 92; DB 4; Length 16;

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052A-54

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RQIKWFOQRNMKWK 16
      |||
Db      1 RQIKWFOQRNMKWK 16
      |||

RESULT 25
US-09-610-220B-9
; Sequence 9, Application US/08610220B
; Patent No. 6635738
; GENERAL INFORMATION:
; APPLICANT: Troy, Carol M.
; TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/610,220B
; FILING DATE: MAR-04-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 48332/JFW/JML
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-09-610-220B-9

Query Match      100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RQIKWFOQRNMKWK 16
      |||
Db      1 RQIKWFOQRNMKWK 16
      |||

RESULT 24
US-09-775-052A-54
; Sequence 54, Application US/09775052A
; Patent No. 6645501
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57

Search completed: May 24, 2004, 17:29:49
Job time : 12.4595 secs
```


GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: May 24, 2004, 17:19:01 ; Search time 30.0541 Seconds
 (without alignments)
 148.491 Million cell updates/sec

Title: US-09-977-349-1
 Perfect score: 92
 Sequence: 1 RQIKWFWQRNMKWK 16

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 100 summaries

Database : Published Applications AA:
 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 15: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		SUMMARIES			
Result No.	Score	Query Match	Length DB ID	Description	
1	92	100.0	16	8 US-08-610-220A-9	Sequence 9, Appli
2	92	100.0	16	9 US-09-214-371-43	Sequence 43, Appl
3	92	100.0	16	9 US-09-780-070-38	Sequence 38, Appl
4	92	100.0	16	9 US-09-150-623-9	Sequence 9, Appli
5	92	100.0	16	9 US-09-731-023A-10	Sequence 10, Appl
6	92	100.0	16	9 US-09-854-204-1	Sequence 5, Appli
7	92	100.0	16	9 US-09-900-147-8	Sequence 1, Appli
8	92	100.0	16	9 US-09-792-480-29	Sequence 8, Appli
9	92	100.0	16	9 US-09-785-802A-2	Sequence 29, Appl
10	92	100.0	16	9 US-09-785-802A-5	Sequence 2, Appli
11	92	100.0	16	9 US-09-902-432-32	Sequence 5, Appli
12	92	100.0	16	9 US-09-953-031A-10	Sequence 32, Appl
13	92	100.0	16	9 US-09-981-286A-3	Sequence 10, Appl
14	92	100.0	16	10 US-09-962-967A-6	Sequence 3, Appli
15	92	100.0	16	10 US-09-912-414-6	Sequence 6, Appli

16	92	100.0	16	10 US-09-775-052-54	Sequence 54, Appl
17	92	100.0	16	10 US-09-295-189-4	Sequence 4, Appli
18	92	100.0	16	10 US-09-965-876A-1	Sequence 1, Appli
19	92	100.0	16	11 US-09-933-780C-2	Sequence 2, Appli
20	92	100.0	16	12 US-10-286-964-3	Sequence 3, Appli
21	92	100.0	16	12 US-10-432-291-6	Sequence 6, Appli
22	92	100.0	16	12 US-10-357-826A-2	Sequence 2, Appli
23	92	100.0	16	12 US-10-603-409-12	Sequence 12, Appl
24	92	100.0	16	12 US-09-748-063-3	Sequence 3, Appli
25	92	100.0	16	12 US-09-779-791A-3	Sequence 3, Appli
26	92	100.0	16	12 US-10-083-960-29	Sequence 29, Appl
27	92	100.0	16	12 US-10-148-549-1	Sequence 1, Appli
28	92	100.0	16	12 US-10-136-187-21	Sequence 21, Appl
29	92	100.0	16	12 US-10-261-161-1	Sequence 1, Appli
30	92	100.0	16	12 US-10-261-161-1	Sequence 26, Appl
31	92	100.0	16	13 US-10-024-935-12	Sequence 12, Appl
32	92	100.0	16	13 US-10-007-363-3	Sequence 3, Appli
33	92	100.0	16	14 US-10-071-512A-2	Sequence 2, Appli
34	92	100.0	16	14 US-10-239-804-3	Sequence 3, Appli
35	92	100.0	16	14 US-10-077-555-3	Sequence 3, Appli
36	92	100.0	16	14 US-10-209-421-29	Sequence 29, Appl
37	92	100.0	16	14 US-10-229-915-2	Sequence 2, Appli
38	92	100.0	16	14 US-10-185-084-3	Sequence 3, Appli
39	92	100.0	16	14 US-10-252-012-5	Sequence 5, Appli
40	92	100.0	16	14 US-10-075-869-19	Sequence 19, Appl
41	92	100.0	16	14 US-10-013-815-19	Sequence 19, Appl
42	92	100.0	16	14 US-10-136-738-10	Sequence 10, Appl
43	92	100.0	16	14 US-10-210-660-1	Sequence 1, Appli
44	92	100.0	16	14 US-10-210-660-25	Sequence 25, Appl
45	92	100.0	16	14 US-10-156-570A-21	Sequence 21, Appl
46	92	100.0	16	14 US-10-201-394A-14	Sequence 14, Appl
47	92	100.0	16	14 US-10-017-672-11	Sequence 11, Appl
48	92	100.0	16	14 US-10-201-389A-14	Sequence 14, Appl
49	92	100.0	16	14 US-10-161-051-1	Sequence 1, Appli
50	92	100.0	16	14 US-10-358-365-10	Sequence 10, Appl
51	92	100.0	16	14 US-10-061-607A-2	Sequence 2, Appli
52	92	100.0	16	14 US-10-405-339-44	Sequence 44, Appl
53	92	100.0	16	15 US-10-366-493-19	Sequence 19, Appl
54	92	100.0	16	15 US-10-444-662-2	Sequence 2, Appli
55	92	100.0	16	15 US-10-185-593-2	Sequence 2, Appli
56	92	100.0	16	15 US-10-413-160-38	Sequence 38, Appl
57	92	100.0	16	15 US-10-462-138-10	Sequence 10, Appl
58	92	100.0	16	15 US-10-369-226-47	Sequence 47, Appl
59	92	100.0	16	15 US-10-353-678-2	Sequence 2, Appli
60	92	100.0	16	15 US-10-357-529-8	Sequence 8, Appli
61	92	100.0	17	9 US-09-854-204-19	Sequence 19, Appl
62	92	100.0	17	9 US-09-785-802A-3	Sequence 3, Appli
63	92	100.0	17	11 US-09-933-780C-3	Sequence 3, Appli
64	92	100.0	17	12 US-10-602-303-3	Sequence 8, Appli
65	92	100.0	17	12 US-10-421-548-8	Sequence 8, Appli
66	92	100.0	17	13 US-10-007-761-8	Sequence 30, Appl
67	92	100.0	17	14 US-10-209-421-30	Sequence 1, Appli
68	92	100.0	17	14 US-10-229-915-1	Sequence 17, Appl
69	92	100.0	17	14 US-10-210-660-20	Sequence 20, Appl
70	92	100.0	17	14 US-10-210-660-22	Sequence 22, Appl
71	92	100.0	17	14 US-10-210-660-27	Sequence 27, Appl
72	92	100.0	17	15 US-10-372-003A-29	Sequence 29, Appl
73	92	100.0	17	15 US-10-428-280-15	Sequence 15, Appl
74	92	100.0	17	15 US-10-421-503-66	Sequence 66, Appl
75	92	100.0	18	9 US-09-785-802A-14	Sequence 14, Appl
76	92	100.0	18	9 US-09-949-474-7	Sequence 7, Appli
77	92	100.0	19	14 US-10-118-079-45	Sequence 45, Appl
78	92	100.0	19	14 US-10-210-660-23	Sequence 23, Appl
79	92	100.0	19	15 US-10-407-449-20	Sequence 20, Appl
80	92	100.0	20	9 US-09-854-204-63	Sequence 63, Appl
81	92	100.0	20	9 US-09-949-474-8	Sequence 8, Appli
82	92	100.0	20	14 US-10-210-660-16	Sequence 16, Appl
83	92	100.0	20	14 US-10-210-660-18	Sequence 18, Appl
84	92	100.0	20	14 US-10-210-660-30	Sequence 30, Appl
85	92	100.0	20	14 US-08-610-220A-11	Sequence 11, Appl
86	92	100.0	21	8 US-09-150-623-11	Sequence 11, Appl
87	92	100.0	21	9 US-08-610-220A-10	Sequence 10, Appl
88	92	100.0	22	8 US-08-610-220A-10	

Query Match	100.0%;	Score 92;	DB 9;	Length 16;
Best Local Similarity	100.0%;	Pred. No. 2.6e-06;		

QY 1 RQIKWIFQNRMMKWK 16
| | | | | | | | | |
Db 1 RQIKWIFQNRMMKWK 16

RESULT 8
US-09-792-480-29
; Sequence 29, Application US/09792480
; Patent No. US20020127198A1
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; FILE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000230US
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US/09/792,480
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Antennapedia
; OTHER INFORMATION: homeodomain, Antennapedia-43-58
US-09-792-480-29

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
| | | | | | | | | |
Db 1 RQIKWIFQNRMMKWK 16

RESULT 9
US-09-785-802A-2
; Sequence 2, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Membrane translocation sequence from Penetratin
US-09-785-802A-2

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
| | | | | | | | | |
Db 1 RQIKWIFQNRMMKWK 16

RESULT 10
US-09-785-802A-5
; Sequence 5, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-785-802A-5

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
| | | | | | | | | |
Db 1 RQIKWIFQNRMMKWK 16

RESULT 11
US-09-902-432-32
; Sequence 32, Application US/09902432
; Patent No. US20020160002A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; APPLICANT: Susan G. Jaken
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A30558-A-FWC-A 070156.0597
; CURRENT APPLICATION NUMBER: US/09/902,432
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Penetratin peptide
US-09-902-432-32

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMMKWK 16
| | | | | | | | | |
Db 1 RQIKWIFQNRMMKWK 16

RESULT 12

US-09-953-031A-10
; Sequence 10, Application US/09953031A
; Patent No. US20020177177A1
; GENERAL INFORMATION:
; APPLICANT: Bernards, Rene
; APPLICANT: Zwijssen, Renate
; TITLE OF INVENTION: Interaction Between Cyclin D1 and Steroid Receptor
; TITLE OF INVENTION: Co-Activators and Uses Thereof in Assays
; FILE REFERENCE: 4238/80713
; CURRENT APPLICATION NUMBER: US/09/953,031A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 09/302,305
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Translocation peptide derived from antennapedia
; OTHER INFORMATION: homeodomain protein
US-09-953-031A-10

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 1 RQIKWIFQNRMRMKWK 16

RESULT 13

US-09-981-286A-3
; Sequence 3, Application US/09981286A
; Publication No. US20020192799A1
; GENERAL INFORMATION:
; APPLICANT: Watowich, Stanley J.
; APPLICANT: Weaver, Scott C.
; APPLICANT: Davey, Robert A.
; TITLE OF INVENTION: Drug Discovery Methods
; FILE REFERENCE: 265.00260101
; CURRENT APPLICATION NUMBER: US/09/981,286A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/240,187
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-3

Query Match 100.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 1 RQIKWIFQNRMRMKWK 16

RESULT 14

US-09-962-967A-6
; Sequence 6, Application US/09962967A
; Publication No. US20030004112A1
; GENERAL INFORMATION:

; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
; FILE REFERENCE: 00398-140002
; CURRENT APPLICATION NUMBER: US/09/962,967A
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/441,416
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: US 08/964,302
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-962-967A-6

Query Match 100.0%; Score 92; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 1 RQIKWIFQNRMRMKWK 16

RESULT 15

US-09-912-414-6
; Sequence 6, Application US/09912414
; Publication No. US20030013169A1
; GENERAL INFORMATION:
; APPLICANT: Muller, Rolf
; APPLICANT: Kontermann, Roland E
; APPLICANT: Montigiani, Silvia
; TITLE OF INVENTION: Transcription factor E2F DNA-binding domain inhibitor
; TITLE OF INVENTION: peptides and their use
; FILE REFERENCE: 620-151
; CURRENT APPLICATION NUMBER: US/09/912,414
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: PCT/GB00/00227
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: GB 9901710.5
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-912-414-6

Query Match 100.0%; Score 92; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWIFQNRMRMKWK 16
Db 1 RQIKWIFQNRMRMKWK 16

RESULT 16

US-09-775-052-54
; Sequence 54, Application US/09775052
; Publication No. US20030054000A1
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 09/208,966
; PRIOR FILING DATE: 1998-12-10

```
; PRIOR APPLICATION NUMBER: 60/082,402
; PRIOR FILING DATE: 1998-04-20
; PRIOR APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052-54

Query Match      100.0%; Score 92; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRMRMVKKK 16
   |||||
Db 1 RQIKWFWQRMRMVKKK 16

RESULT 17
US-09-295-189-4
; Sequence 4, Application US/09295189
; Publication No. US20030083273A1
; GENERAL INFORMATION:
; APPLICANT: Woolf, Tod M.
; TITLE OF INVENTION: Improved Antisense Oligomers
; FILE REFERENCE: SRI-004
; CURRENT APPLICATION NUMBER: US/09/295,189
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-295-189-4

Query Match      100.0%; Score 92; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRMRMVKKK 16
   |||||
Db 1 RQIKWFWQRMRMVKKK 16

RESULT 18
US-09-965-876A-1
; Sequence 1, Application US/09965876A
; Publication No. US2003096243A1
; GENERAL INFORMATION:
; APPLICANT: Cellomics, Inc.
; APPLICANT: Busa, William B
; TITLE OF INVENTION: Methods and Reagents for Live-cell Gene Expression Quantification
; FILE REFERENCE: 00-789-A
; CURRENT APPLICATION NUMBER: US/09/965,876A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,407
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-965-876A-1
```

```
Query Match      100.0%; Score 92; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRMRMVKKK 16
   |||||
Db 1 RQIKWFWQRMRMVKKK 16

RESULT 19
US-09-933-780C-2
; Sequence 2, Application US/09933780C
; Publication No. US20030229202A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS INC.
; APPLICANT: GUO, Yong
; APPLICANT: MORSE, Clarence C
; APPLICANT: YAO, Zhengbin
; TITLE OF INVENTION: MEMBRANE PENETRATING PEPTIDES AND USES THEREOF
; FILE REFERENCE: HMR2053 PCT
; CURRENT APPLICATION NUMBER: US/09/933,780C
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,647
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: GB 0103110.3
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Signal sequence peptide from Antennapedia homeodomain
US-09-933-780C-2

Query Match      100.0%; Score 92; DB 11; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQIKWFWQRMRMVKKK 16
   |||||
Db 1 RQIKWFWQRMRMVKKK 16

RESULT 20
US-10-286-964-3
; Sequence 3, Application US/10286964
; Publication No. US20040029791A1
; GENERAL INFORMATION:
; APPLICANT: Fahraeus, Robin
; APPLICANT: Lane, David P.
; TITLE OF INVENTION: Cyclin Dependent Kinase Binding Compounds
; FILE REFERENCE: CCI-003US
; CURRENT APPLICATION NUMBER: US/10/286,964
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/043,560
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-286-964-3

Query Match      100.0%; Score 92; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Query Match	100.0%	Score 92	DB 12	Length 16
Best Local Similarity	100.0%	Pred. No. 2.6e-06		
Matches 15	Conservative	0	Mismatches 0	Indels 0
				Gaps 0

QY 1 ROIKIWFQNRMMKKK 16
| | | | | | | | | | | | | | | | | |
Db 1 ROIKIWFQNRMMKKK 16

RESULT 25

US-09-779-791A-3
; Sequence 3, Application US/09779791A
; Publication No. US2001004417A1
; GENERAL INFORMATION:
; APPLICANT: Mirus Corporation
; APPLICANT: Wolff, Jon A
; APPLICANT: Monahan, Sean D
; APPLICANT: Budker, Vladimir G
; APPLICANT: Slattum, Paul M
; APPLICANT: Rozema, David B
; TITLE OF INVENTION: A Compound Containing a Labile Disulfide Bond
; FILE REFERENCE: Mirus.006.03
; CURRENT APPLICATION NUMBER: US/09/779,791A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 09/312,351
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-779-791A-3

Query Match 100.0%; Score 92; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ROIKIWFQNRMMKKK 16
| | | | | | | | | | | | | | | | | |
Db 1 ROIKIWFQNRMMKKK 16

Search completed: May 24, 2004, 17:28:44
Job time : 32.0541 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:13:28 ; Search time 45.4865 Seconds

(without alignments)
105.599 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IROFKIWFNNRRKFWKX 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004as.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	17	3	AAY93955 Peptide u
2	103	100.0	19	3	AAY93957 Peptide u
3	99	96.1	16	2	AAW33412 Peptide 4
4	99	96.1	16	5	ABB83154 Transduct
5	99	96.1	16	6	AAE33893 Drosophil
6	99	96.1	16	6	ABP70226 Membrane
7	99	96.1	26	4	AAW72140 Anti-alle
8	99	96.1	26	4	AAW72143 Anti-alle
9	99	96.1	27	5	ABB83155 Transduct
10	85	82.5	17	6	ABU09583 Cell perm
11	82.5	80.1	17	3	AAW83575 Modified
12	79	76.7	17	4	AAW85051 N-termina
13	75	72.8	16	2	AAW33411 Peptide 4
14	75	72.8	16	3	AAW83574 Modified
15	75	72.8	16	3	AAW51213 Antennape
16	75	72.8	16	4	AAE12484 Membrane
17	75	72.8	16	6	AAE35567 Antennape
18	75	72.8	16	6	ABU09582 Cell perm
19	75	72.8	16	6	AAE33892 Drosophil
20	75	72.8	16	6	ABB82913 Pro 50 pe
21	75	72.8	16	6	AAO23274 Antennape
22	75	72.8	17	3	AAE13427 Synthetic
23	75	72.8	18	3	AAE13432 Synthetic
24	75	72.8	19	3	AAE13428 Synthetic
25	75	72.8	20	3	AAE13431 Synthetic

26	75	72.8	21	3	AAE13430 Synthetic
27	75	72.8	22	3	AAE13429 Synthetic
28	75	72.8	27	4	AAE12483 Membrane
29	75	72.8	35	6	AAE35578 TA peptid
30	74	71.8	16	4	AAW85059 N-termina
31	70	68.0	22	2	AAW27475 Human TSP
32	69	67.0	17	3	AAW83558 Peptide f
33	68	66.0	17	5	AAU77232 Synthetic
34	68	66.0	26	3	AAE03930 Beta-cate
35	68	66.0	36	3	AAW71011 Penetrati
36	68	66.0	36	3	AAW71011 Penetrati
37	68	66.0	106	3	AAE21030 Human nuc
38	68	66.0	115	3	AAE53629 Human col
39	68	66.0	220	4	AAW41487 Human pol
40	68	66.0	236	4	AAW39701 Human pol
41	68	66.0	243	2	AAW48885 Amino aci
42	67	65.0	16	2	AAW45974 Cysteine
43	67	65.0	16	2	AAW33407 Peptide 4
44	67	65.0	16	2	AAW33410 D-form pe
45	67	65.0	16	2	AAW82958 Oestrogen
46	67	65.0	16	2	AAW56397 Preferred
47	67	65.0	16	2	AAW71270 Antennape
48	67	65.0	16	2	AAW71316 Antennape
49	67	65.0	16	2	AAW30508 Drosophil
50	67	65.0	16	2	AAW91046 Internall
51	67	65.0	16	2	AAW52102 Peptide f
52	67	65.0	16	2	AAW08859 Peptide p
53	67	65.0	16	2	AAW13509 Signal se
54	67	65.0	16	3	AAW87920 Drosophil
55	67	65.0	16	3	AAW27060 Beta-cate
56	67	65.0	16	3	AAW93667 Peptide w
57	67	65.0	16	3	AAW67966 Carboxyl
58	67	65.0	16	3	AAW93551 Amino aci
59	67	65.0	16	3	AAW55818 Signal se
60	67	65.0	16	3	AAW71008 Drosophil
61	67	65.0	16	3	AAW83572 Modified
62	67	65.0	16	3	AAW51212 Antennape
63	67	65.0	16	3	AAW51167 Drosophil
64	67	65.0	16	3	AAW10343 Peptide A
65	67	65.0	16	3	AAE19251 Fragment
66	67	65.0	16	3	AAW93178 Protegrin
67	67	65.0	16	3	AAW35694 Peptide a
68	67	65.0	16	3	AAW22025 Membrane
69	67	65.0	16	3	AAW29423 ANTP pept
70	67	65.0	16	3	AAW03927 Internali
71	67	65.0	16	3	AAW93954 Peptide u
72	67	65.0	16	3	AAW29574 Antennape
73	67	65.0	16	3	AAE14785 Carrier m
74	67	65.0	16	3	AAE14761 Drosophil
75	67	65.0	16	4	AAW73091 Rheumatol
76	67	65.0	16	4	AAW60004 Internali
77	67	65.0	16	4	AAW70753 Cell memb
78	67	65.0	16	4	AAW02974 Protein t
79	67	65.0	16	4	AAW60671 Antennape
80	67	65.0	16	4	AAU06064 Drosophil
81	67	65.0	16	4	AAW49914 HIF-lalph
82	67	65.0	16	4	AAW66996 Antennape
83	67	65.0	16	4	AAU00813 Fruit fly
84	67	65.0	16	4	AAE12205 Membrane
85	67	65.0	16	5	ABB78030 Peptide d
86	67	65.0	16	5	ABG78985 Cell pene
87	67	65.0	16	5	AAW47331 Penetrati
88	67	65.0	16	5	AAE15616 Drosophil
89	67	65.0	16	5	AAO15781 Antennape
90	67	65.0	16	5	AAU78345 Antennape
91	67	65.0	16	5	ABG68406 Transloca
92	67	65.0	16	5	ABB83153 Transduct
93	67	65.0	16	5	ABP53809 Penetrati
94	67	65.0	16	5	AAE25451 Drosophil
95	67	65.0	16	5	ABG75500 Fruit fly
96	67	65.0	16	5	ABG60447 Selective
97	67	65.0	16	5	ABG30763 E2F-1/cyc
98	67	65.0	16	5	ABG30763

Tue May 25 09:20:17 2004

us-09-977-349-2.rag

99 67 65.0 16 5 AAU78912 Aau78912 Antennape
100 67 65.0 16 5 ABB81174 ABB81174 Antennape

ALIGNMENTS

RESULT 1
AAU93955
ID AAU93955 standard; peptide; 17 AA.

XX AC AAY93955;
XX DT 03-OCT-2000 (first entry)

XX Peptide used to construct inhibitory pseudopeptide compounds.

XX Pseudopeptide compound; inhibitor; tyrosine kinase;
KW growth factor receptor binding protein; Grb2; proliferative disorder;
KW cancer; metastasis; antennapaedia protein.

XX Synthetic.

XX WO200039153-A1.

XX 06-JUL-2000.

XX 24-DEC-1999; 99WO-FR003289.

XX 24-DEC-1998; 98FR-00016459.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Garbay C, Liu W, Vidal M, Roques BP;

XX WPI; 2000-475683/41.

XX Pseudopeptides, useful for the treatment of proliferative process
PT disorders, cancers, and metastases, inhibit paths activated by proteins
PT with a tyrosine kinase activity.

XX Claim 1; Page 32; 43pp; French.

XX The specification describes pseudopeptide compounds. These compounds are
CC inhibitors of routes activated by proteins having a tyrosine kinase
CC activity. The compounds also have high affinity for a growth factor
CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
CC the treatment of disorders of proliferative processes, cancers, and
CC metastases. The present sequence is derived from the antennapaedia
CC protein, and is a fragment of the pseudopeptide compounds of the
CC invention

XX Sequence 17 AA;

Query Match 100.0%; Score 103; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IROPKIWFNNRRKPWK 17
DB 1 IROPKIWFNNRRKPWK 17

RESULT 2

AAU93957
ID AAU93957 standard; peptide; 19 AA.

XX AC AAY93957;

XX 03-OCT-2000 (first entry)

XX Peptide used to construct inhibitory pseudopeptide compounds.

XX Pseudopeptide compound; inhibitor; tyrosine kinase;
KW growth factor receptor binding protein; Grb2; proliferative disorder;
KW cancer; metastasis; antennapaedia protein.

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..2

XX WO200039153-A1.

XX 06-JUL-2000.

XX 24-DEC-1999; 99WO-FR003289.

XX 24-DEC-1998; 98FR-00016459.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Garbay C, Liu W, Vidal M, Roques BP;

XX WPI; 2000-475683/41.

XX Pseudopeptides, useful for the treatment of proliferative process
PT disorders, cancers, and metastases, inhibit paths activated by proteins
PT with a tyrosine kinase activity.

XX Claim 1; Page 32; 43pp; French.

XX The specification describes pseudopeptide compounds. These compounds are
CC inhibitors of routes activated by proteins having a tyrosine kinase
CC activity. The compounds also have high affinity for a growth factor
CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
CC the treatment of disorders of proliferative processes, cancers, and
CC metastases. The present sequence is derived from the antennapaedia
CC protein, and is a fragment of the pseudopeptide compounds of the
CC invention

XX Sequence 19 AA;

Query Match 100.0%; Score 103; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IROPKIWFNNRRKPWK 17
DB 3 IROPKIWFNNRRKPWK 19

RESULT 3

AAW33412
ID AAW33412 standard; peptide; 16 AA.

XX AC AAW33412;

XX 17-MAR-1998 (first entry)

XX Peptide 43-58 (3Pro) of homeodomain Antp.

XX homeodomain; transcription factor; Antennapedia; Antp; vector;

KW transfection; hydrophobic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "in determining the ability of this sequence to be
FT internalised in cells, a biotin-aminopentanoyl group was
FT attached to the N-terminal"

XX WO9712912-A1.

CC cell anaemia, a viral disease, AIDS, a bacterial disease, herpes,
 CC pneumonia, tuberculosis or an inflammatory disease, in a patient. It is
 CC also used in gene therapy. The present sequence is *Drosophila* sp.
 CC antennapedia third helix peptide used to illustrate the method of the
 CC invention
 XX
 SQ Sequence 16 AA;

Query Match 96.1%; Score 99; DB 6; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RQKIWFNRRKPKWK 17
 |||||
 DB 1 RQKIWFNRRKPKWK 16

RESULT 6
 ABP70226
 ID ABP70226 standard; peptide; 16 AA.
 XX
 AC ABP70226;
 XX
 DT 07-APR-2003 (first entry)
 XX
 DE Membrane translocating peptide from homeobox domain penetration region.
 XX
 KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;
 KW cancer; infection; immune deficiency; gene defect; genetic disease;
 KW membrane translocating peptide.
 XX

OS Unidentified.
 XX WO20028318-A2.
 XX 07-NOV-2002.
 XX 30-APR-2002; 2002WO-US013609.
 XX 30-APR-2001; 2001US-0287786P.
 XX (TARG-) TARGETED GENETICS CORP.
 PA (EMER-) EMERALD GENE SYSTEMS LTD.
 XX
 PI Harvie P, Paul R, Cudmore S, O'mahony DJ;
 XX WPI; 2003-183837/18.
 XX

PT Lipid-nucleic acid complex useful for delivering a nucleic acid to a
 PT cell, comprises compacted nucleic acid, polycation, targeting factor and
 PT lipid, and does not comprise protamine or its salt.
 XX

PS Disclosure; Page 42; 259pp; English.
 XX
 CC The specification describes a lipid-nucleic acid complex, comprising a
 CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but
 CC not a protamine. The targeting factor increases cellular bioavailability
 CC of the nucleic acid without interaction with a specific outer cell
 CC surface membrane receptor. The mean diameter of the complex is greater
 CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is
 CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene
 CC therapy. It reduces levels of inflammatory cytokines such as tumour
 CC necrosis factor- α . The complex is useful for manufacturing a
 CC medicament for treating or diagnosing a variety of diseases, conditions
 CC or syndromes such as cancer, bacterial, viral or parasitic infections,
 CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited
 CC genetic diseases). The present sequence represents a membrane
 CC translocating peptide, which is used as the targeting factor in lipid-
 CC nucleic acid complexes of the invention
 XX

SQ Sequence 16 AA;

Query Match 96.1%; Score 99; DB 6; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 RQKIWFNRRKPKWK 17
 |||||
 DB 1 RQKIWFNRRKPKWK 16

RESULT 7
 AAY72140
 ID AAY72140 standard; peptide; 26 AA.
 XX

AC AAY72140;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Anti-allergic peptide 3.

XX Fruit fly; anti-allergic peptide; therapeutic; migraine; psoriasis;
 KW multiple sclerosis; nasal allergy; mast cell degranulation; histamine;
 KW allergy; eye; skin; acute urticaria; interstitial cystitis; vasotropic;
 KW psychogenic; bowel disease; dermatological; anti-inflammatory; G alpha13;
 KW neuroprotective; antipsoriatic; fusion peptide; exocytosis; asthma;
 KW *Drosophila* transcription factor.

OS *Drosophila* sp.
 OS Unidentified.
 OS Chimeric.

FH Key Location/Qualifiers
 FT Peptide 1..16
 FT /label= Signal peptide
 FT /note= "Signal sequence of homeodomain of *Drosophila*
 FT transcription factor"
 FT 17..26
 FT /label= G_alpha13 peptide
 FT /note= "Corresponds to C-terminal sequence of G_alpha13"

XX WO200078346-A1.

XX 28-DEC-2000.

XX 14-JUN-2000; 2000WO-IL000346.

XX 17-JUN-1999; 99IL-00130526.

XX (ALLE-) ALLERGENE LTD.

XX Eisenberg R, Raz T;

XX WPI; 2001-080758/09.

XX Novel anti-allergic agents for treating allergic conditions such as
 PT allergic reactions in eye, skin, nasal allergy, asthma, migraines, has
 PT peptides for cell penetration and reducing mast cell degranulation.

XX Example 1; Page 13; 63pp; English.

XX The present sequence is anti-allergic peptide 3 consisting of a signal
 CC sequence of homeodomain of *Drosophila* transcription factor, linked to the
 CC C-terminal G_alpha13 sequence. This C-terminal G_alpha13 appears to
 CC mediate the peptidergic pathway leading to exocytosis in mast cells. The
 CC invention relates to therapeutic complex molecules which are useful as
 CC anti-allergic agents. These anti-allergic agents are useful for treating
 CC allergic conditions such as nasal allergy, allergic reaction in the eye
 CC or skin, acute urticaria, psoriasis, psychogenic or allergic asthma,
 CC interstitial cystitis, bowel diseases, migraines and multiple sclerosis.
 CC The therapeutic complex is highly specific, direct and provides targeted
 CC treatment of allergies and related inflammatory conditions. It comprises
 CC molecules having at least a first segment i.e., a signal peptide which is
 CC competent for the importation of the complex into the mast cells, and a
 CC second segment which is having the anti-allergic effect is able to block
 CC or significantly reduce the G protein-mediated contribution to mast cell

CC related inflammatory conditions. It comprises molecules having at least a
 CC first segment ie., a signal peptide which is competent for the
 CC importation of the complex into the mast cells, and a second segment
 CC which is having the anti-allergic effect is able to block or
 CC significantly reduce the G protein-mediated contribution to mast cell
 CC degranulation and in turn the release of histamine. The invention also
 CC discloses methods for preventing and treating allergies

XX SQ Sequence 26 AA;

Query Match 96.1%; Score 99; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17
 |||||
 Db 1 RQPKWFPNRRKPWK 16

RESULT 9
 ABB83155
 ID ABB83155 standard; peptide; 27 AA.

XX AC ABB83155;
 XX DT 08-AUG-2002 (first entry)
 XX DE Transduction domain/nuclear export sequence peptide NES-Pen3P.
 XX KW Transduction domain; nuclear export; tight epithelial junction;
 XX KW blood-brain barrier; choroid plexus.
 XX OS Unidentified.
 XX PH Key Location/Qualifiers
 FT Modified-site 1 /note= "Biotinylated Gln"
 FT FT
 XX WO200239947-A2.
 XX PD 23-MAY-2002.
 XX PF 20-NOV-2001; 2001WO-FR003631.
 XX PR 20-NOV-2000; 2000FR-00014945.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Joliot A, Dupont E, Prochiantz A;
 XX DR WPI; 2002-471597/50.
 XX PT Use of vector a peptide containing transduction domain and nuclear export
 XX PT sequence for transporting diagnostic or therapeutic agents across tight
 XX PT epithelial junctions.
 XX PS Example 4; Page 16; 30pp; French.
 XX CC The present invention relates to peptides that include at least one
 CC transduction domain (TD) and a nuclear export sequence (NES) to prepare a
 CC vector for transporting diagnostic or therapeutic agents across a tight
 CC epithelial junction. The present sequence is one such peptide. The
 CC combination of TD and NES increases the efficiency of transport through
 CC tight junctions, compared with use of TD alone. The vectors are useful
 CC for delivering nucleic acids, polypeptides, peptide nucleic acids, or a
 CC nucleotide analog (e.g. an antineoplastic agent) across the blood-brain
 CC barrier or the choroid plexus

XX SQ Sequence 27 AA;
 Query Match 96.1%; Score 99; DB 5; Length 27;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC degranulation and in turn the release of histamine. The invention also
 CC discloses methods for preventing and treating allergies

XX SQ Sequence 26 AA;

Query Match 96.1%; Score 99; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17
 |||||
 Db 1 RQPKWFPNRRKPWK 16

RESULT 8
 AAY72143
 ID AAY72143 standard; peptide; 26 AA.

XX AC AAY72143;
 XX DT 24-APR-2001 (first entry)
 XX DE Anti-allergic peptide 6.
 XX KW Fruit fly; anti-allergic peptide; therapeutic; migraine; psoriasis;
 KW multiple sclerosis; nasal allergy; mast cell degranulation; histamine;
 KW allergy; eye; skin; acute urticaria; interstitial cystitis; vasotropic;
 KW psychogenic; bowel disease; dermatological; antiinflammatory; G alpat;
 KW neuroprotective; antiporiatic; fusion peptide; exocytosis; asthma;
 KW Drosophila transcription factor.
 XX OS Drosophila sp.
 OS Unidentified.
 OS Chimeric.
 XX PH Key Location/Qualifiers
 FT Peptide 1..16
 FT /label= Signal peptide
 FT /note= "Signal sequence of homeodomain of Drosophila
 FT transcription factor"
 FT 17..26
 FT /label= G alpat peptide
 FT /note= "Corresponds to C-terminal sequence of G alpat"
 XX WO200078346-A1.
 XX PD 28-DEC-2000.
 XX PF 14-JUN-2000; 2000WO-IL000346.
 XX PR 17-JUN-1999; 99IL-00130526.
 XX PA (ALLE-) ALLERGENE LTD.
 XX PI Eisenberg R, Raz T;
 XX DR WPI; 2001-080758/09.
 XX CC Novel anti-allergic agents for treating allergic conditions such as
 XX PT allergic reactions in eye, skin, nasal allergy, asthma, migraines, has
 XX PT peptides for cell penetration and reducing mast cell degranulation.
 XX PS Example 1; Page 13; 63pp; English.
 XX CC The present sequence is anti-allergic peptide 6 consisting of a signal
 CC sequence of homeodomain of Drosophila transcription factor, linked to the
 CC C-terminal G alpat sequence. The invention relates to therapeutic
 CC complex molecules which are useful as anti-allergic agents. These anti-
 CC allergic agents are useful for treating allergic conditions such as nasal
 CC allergy, allergic reaction in the eye or skin, acute urticaria,
 CC psoriasis, psychogenic or allergic asthma, interstitial cystitis, bowel
 CC diseases, migraines and multiple sclerosis. The therapeutic complex is
 CC highly specific, direct and provides targeted treatment of allergies and

QY 2 ROPKIWFNRRKPWK 17
 12 ROPKIWFNRRKPWK 27
 Db
 RESULT 10
 ABU09583
 ID ABU09583 standard; peptide; 16 AA.
 XX
 AC ABU09583;
 XX
 DT 09-JUL-2003 (first entry)
 XX
 DE Cell permeant peptide #4.
 XX
 KW Capsid protein; CCD; apteain library; cell permeant peptide; cell death;
 KW pathogen; viral infection; bacterial infection; toxin; Rickets;
 KW fungal infection; mortality; morbidity; biological warfare.
 XX
 OS Unidentified.
 XX
 PN US2002192799-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 15-OCT-2001; 2001US-00981286.
 XX
 PR 13-OCT-2000; 2000US-0240187P.
 PR
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA
 PI Watowich SJ, Weaver SC, Davey RA;
 XX
 DR WPI; 2003-428794/40.
 XX
 PT Novel collection of polypeptides which comprise fragment of Venezuelan
 PT equine encephalitis virus capsid polypeptide carboxy terminal portions,
 PT useful for displaying variable amino acid sequences on the surface of the
 PT polypeptide.
 XX
 PS Claim 3; Page 12; 25pp; English.
 XX
 CC The invention relates to a collection of polypeptides comprising two
 CC polypeptides (referred to as an apteain library) each of which has a
 CC fragment of amino acids 119-275 of Venezuelan equine encephalitis virus
 CC capsid polypeptide carboxy terminal portion (ABU09576, also known as
 CC CCD), beginning at any of amino acids 119-124 and ending at any of amino
 CC acids 258-275, where at least two consecutive amino acids within amino
 CC acids 129-137, 182-189, 257-264 of CCD are replaced by a variant amino
 CC acid sequence. Also included are a population of cells comprising two or
 CC more cells (where each member of the population comprises one polypeptide
 CC of an apteain), a fusion protein chosen from an apteain fused to a cell
 CC permeant peptide, a cell comprising the fusion protein, a collection of
 CC polynucleotides (comprising at least two polynucleotides, each
 CC comprising an apteain), a vector comprising a member of the
 CC polynucleotide collection, a population of cells comprising two or more
 CC cells containing members of the collection of fusion proteins and
 CC crystallising a polypeptide comprising CCD. The apteains are useful for
 CC identifying a polypeptide within a collection that prevents cell death
 CC after exposure to a pathogen (e.g. a virus or microbe such as bacterium,
 CC Rickettsia or fungus), or a toxin such as a biological toxin or chemical
 CC toxin and is useful for identifying a polypeptide within a collection
 CC that binds to a pathogen, toxin, polypeptide or a polynucleotide. The
 CC apteain is useful in discovery of drugs that can be used to prevent
 CC mortality and morbidity associated with pathogens and toxins, including
 CC high lethal agents that could be used in biological warfare. The present
 CC sequence represents a cell permeant peptide used to make fusion proteins
 CC with the CCD-based apteains
 XX
 SQ Sequence 16 AA;
 Query Match 80.1%; Score 82.5; DB 3; Length 17;
 Best Local Similarity 88.2%; Pred. No. 3.8e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 2 ROPKIWFNRRKPWK 17
 1 ROPKIWFNRRKPWK 16
 Db
 RESULT 11
 AAY83575
 ID AAY83575 standard; peptide; 17 AA.
 XX
 AC AAY83575;
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Modified peptide 45,50,55 Pro of penetratin.
 XX
 KW Penetratin; translocation; membrane; drug delivery; therapy; treatment;
 KW conjugate; vector.
 XX
 OS Synthetic.
 XX
 PN WO200029427-A2.
 XX
 PD 25-MAY-2000.
 XX
 PF 11-NOV-1999; 99WO-GB003750.
 XX
 PR 13-NOV-1998; 98GB-00025000.
 PR 13-NOV-1998; 98GB-00025001.
 PR 04-FEB-1999; 99GB-00002522.
 PR 04-FEB-1999; 99GB-00002525.
 PR 22-JUN-1999; 99GB-00014578.
 PR
 XX (CYCL-) CYCLACEL LTD.
 XX
 PI Fischer MP, Zhelev N;
 XX
 DR WPI; 2000-387734/33.
 XX
 PT New membrane translocation peptide carrier group for delivering
 PT therapeutic agents into target cells comprises specified sequence of
 PT amino acids.
 XX
 PS Example 4; Page 29; 59pp; English.
 XX
 CC Penetratin is a membrane translocation polypeptide and as such, active
 CC peptide fragments of penetratin can be used to translocate conjugated
 CC molecules across the cell membrane e.g. drugs. 13 chemical entities are
 CC described for use in the method such as paclitaxel-
 CC 2-succinimidopropionyl-CbetaA-RRMKWK-NH₂ and podophyllotoxin-4-
 CC succinimidopropionyl-CbetaA-RRMKWK-NH₂. The method has applications as
 CC a drug delivery system for treatment and therapy. The resulting
 CC conjugated molecules exhibit high immunogenicity, solubility and
 CC clearance. The penetratin peptide fragment may be truncated and or have
 CC amino acid substitutions. See GENESSEQ records AAY83520-Y83581
 XX
 SQ Sequence 17 AA;
 Query Match 80.1%; Score 82.5; DB 3; Length 17;
 Best Local Similarity 88.2%; Pred. No. 3.8e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 2 ROPKIWFNRRKPWK 17
 1 ROPKIWFNRRKPWK 17
 Db
 RESULT 12
 AAB85051
 ID AAB85051 standard; peptide; 17 AA.
 Query Match 80.1%; Score 82.5; DB 3; Length 17;
 Best Local Similarity 88.2%; Pred. No. 3.8e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 2 ROPKIWFNRRKPWK 17
 1 ROPKIWFNRRKPWK 17
 Db

XX 07-FEB-2000; 2000US-0180823P.
 PR 22-FEB-2000; 2000US-0184057P.
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX Brandt C, Bultmann H;
 XX WPI; 2001-638840/73.
 XX Peptides comprising membrane transiting peptides useful for treating or
 PT preventing a virus infection, e.g., human immunodeficiency virus, herpes
 PT simplex virus and cytomegalovirus.
 XX Claim 7; Page 15; 43pp; English.
 XX The invention relates to peptides comprising membrane transiting peptides
 CC with antiviral properties. The peptides are useful for treating or
 CC preventing a virus infection in a warm blooded animal, e.g., enveloped
 CC viruses such as human immunodeficiency virus (HIV), herpes simplex virus
 CC (HSV), cytomegalovirus (CMV) and non-enveloped virus. Preferably, the
 CC peptides are useful for treating or preventing infections from one or
 CC more HSVs. The antiviral peptides are used for treating viral infections
 CC of the skin or part of the oral or genital cavity. The present sequence
 CC is membrane transiting antiviral peptide
 XX
 XX Sequence 16 AA;
 Query Match 72.8%; Score 75; DB 4; Length 16;
 Best Local Similarity 81.2%; Pred. No. 0.00043;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RQPKIWFPPNRKPKWK 17
 DB 1 RQIKIWFPPNRKMKKK 16
 RESULT 17
 ID AAE35567
 AC AAE35567;
 DT 17-JUN-2003 (first entry)
 DE Antennapedia homeodomain (AntpHD) protein translocating peptide.
 XX Fusion agent; immunogenic; proliferative disease; infectious disease;
 KW cancer; therapy; vaccine; melanoma; Trojan antigen; TA.
 XX Antennapedia.
 OS WO200294994-A2.
 XX 28-NOV-2002.
 XX 20-MAY-2002; 2002WO-US015992.
 XX 18-MAY-2001; 2001US-0291874P.
 XX (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
 XX Celis E;
 XX WPI; 2003-140367/13.
 XX Fusion agent useful for preventing and treating an infectious disease, or
 PT a proliferative disease, such as cancer, comprises a transport domain,
 PT two cleavage sites, a peptide epitope and a biologically active agent.
 XX Claim 9; Page 14; 72pp; English.
 XX The invention relates to a fusion agent (Trojan antigen; TA) comprising a

CC transport domain, two cleavage sites, a peptide epitope recognised by an
 CC antigen-specific receptor on an effector T-lymphocyte precursor cell and
 CC a biologically active agent, where there is a cleavage site between the
 CC peptide epitope and the biologically active agent and between each
 CC biologically active agent. The fusion agent is used to make a cell
 CC immunogenic or antigenic. It is also useful for preventing and treating
 CC an infectious disease such as viral, bacterial, protozoal, fungal or
 CC yeast disease, or proliferative disease such as cancer (e.g. melanoma,
 CC neural tissue, gastrointestinal, breast, lung, ovarian, testicular,
 CC prostate, cervical, bladder, vaginal, liver, renal, bone, haematological
 CC or vascular tissue cancer). The invention is used as vaccines. The
 CC present sequence is Antennapedia homeodomain (AntpHD) protein
 CC translocating peptide. This sequence is used in the exemplification of
 CC the invention
 XX
 XX Sequence 16 AA;
 Query Match 72.8%; Score 75; DB 6; Length 16;
 Best Local Similarity 81.2%; Pred. No. 0.00043;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RQPKIWFPPNRKPKWK 17
 DB 1 RQIKIWFPPNRKMKKK 16
 RESULT 18
 ID ABU09582
 XX ABU09582 standard; peptide; 16 AA.
 AC ABU09582;
 XX 09-JUL-2003 (first entry)
 DT Cell permeant peptide #3.
 DE Capsid protein; CCD; adaptin library; cell permeant peptide; cell death;
 KW pathogen; viral infection; bacterial infection; toxin; Rickets;
 KW fungal infection; mortality; morbidity; biological warfare.
 XX Unidentified.
 OS US2002192799-A1.
 XX 19-DEC-2002.
 XX 15-OCT-2001; 2001US-00981286.
 XX 13-OCT-2000; 2000US-0240187P.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Watowich SJ, Weaver SC, Davey RA;
 XX WPI; 2003-428794/40.
 XX Novel collection of polypeptides which comprise fragment of Venezuelan
 PT equine encephalitis virus capsid polypeptide carboxy terminal portions,
 PT useful for displaying variable amino acid sequences on the surface of the
 PT polypeptide.
 XX Claim 3; Page 12; 25pp; English.
 XX The invention relates to a collection of polypeptides comprising two
 CC polypeptides (referred to as an adaptin library) each of which has a
 CC fragment of amino acids 119-275 of Venezuelan equine encephalitis virus
 CC capsid polypeptide carboxy terminal portion (ABU09576, also known as
 CC CCD), beginning at any of amino acids 119-124 and ending at any of amino
 CC acids 258-275, where at least two consecutive amino acids within amino
 CC acids 129-137, 182-189, 257-264 of CCD are replaced by a variant amino
 CC acid sequence. Also included are a population of cells comprising two or
 CC more cells (where each member of the population comprises one polypeptide
 CC of an adaptin), a fusion protein chosen from an adaptin fused to a cell

QY 2 RQKIWFNRRKPKWK 17
 |||||
 Db 1 RQKIWFNRRKPKWK 16

RESULT 21
 AAO23274
 ID AAO23274 standard; peptide; 16 AA.

XX AC AAO23274;

XX DT 25-SEP-2003 (first entry)

DE DE Antennapedia homeodomain peptide to target intracellular polypeptides.

KW Immunoregulation; T cell response; 4-1BB binding agent; CD137; TNF;
 KW tumour necrosis factor; immunogenic; gene therapy; leukaemia; cancer;
 KW astrocytoma; infectious disease; immunostimulant; antennapedia; anergy.

XX OS Unidentified.

PN WO2003049755-A1.

XX 19-JUN-2003.

XX 09-OCT-2002; 2002WO-US032364.

XX 09-OCT-2001; 2001US-0328004P.

XX (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

XX Chen L, Strome SE;

XX WPI; 2003-532879/50.

PT Generating an enhanced immune response in a subject comprises
 PT administering to the subject an immunogenic stimulus (e.g. a tumor-
 PT associated antigen) and an agonistic 4-1BB-binding agent (e.g. antibody
 PT that binds to 4-1BB).

PS Disclosure; Page 21; 81pp; English.

CC This invention relates to a novel method of immunoregulation,
 CC specifically T cell response regulation by enhancing the immune response
 CC with 4-1BB binding agents. 4-1BB (also known as CD137) is a member of the
 CC tumour necrosis factor (TNF) receptor family, and is expressed on the
 CC surface of activated T-cells. Administration of a 4-1BB binding agent and
 CC a tumour specific immunogenic stimulus has been found to result in the
 CC regression of that tumour. The binding agent can be an antibody that
 CC binds 4-1BB or a natural ligand of 4-1BB (or a functional fragment
 CC thereof), whereas the immunogenic stimulus is either a tumour associated
 CC peptide fragment or a molecule produced by an infectious microorganism
 CC e.g. a retrovirus or protozoan parasite. As such, the present invention
 CC describes a method of gene therapy for the treatment of, for example,
 CC leukaemia, astrocytomas, colon cancer, testicular cancer, bone and
 CC vascular tumours. In addition, with an appropriate immunostimulant it can
 CC provide a therapy or prophylaxis for various infectious diseases.
 CC Furthermore, it has been shown that this method prevents anergy (a
 CC depression in immunological function), and can also reverse established
 CC anergy, in CD8+ T-cells. This peptide sequence is the antennapedia
 CC homeodomain peptide that acts to direct proteins, such as the
 CC immunostimulant or 4-1BB binding agents, to specific cellular
 CC compartments, the method of the invention

XX Sequence 16 AA;

Query Match 72.8%; Score 75; DB 6; Length 16;
 Best Local Similarity 81.2%; Pred. No. 0.00043;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPKWK 17
 |||||

Db 1 RQKIWFNRRKPKWK 16

RESULT 22

AAB13427
 ID AAB13427 standard; peptide; 17 AA.

XX AC AAB13427;

XX DT 23-NOV-2000 (first entry)

XX DE Synthetic alpha smooth muscle actin inhibitor # 7.

XX Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.

XX OS Synthetic.

XX Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Tetrapeptide of Acetyl-EEED linked on the epsilon-
 amino group of lysyl residue"

XX WO200038733-A1.
 PN 06-JUL-2000.

XX 15-DEC-1999; 99WO-EP009964.

XX 24-DEC-1998; 98EP-00204396.

XX (UNIO) UCB-BIOPRODUCTS SA.

XX Gabbiani G, Scarso A;

XX WPI; 2000-452308/39.

XX A peptidic product for prevention and treatment of a disease related to
 PT alpha-SM actin expression comprises a tetrapeptide associated with a
 PT chemical entity that is able to introduce the tetrapeptide into the cell.

XX Claim 7; Page 24; 31pp; English.

CC The present invention relates to novel peptides comprising of a specific
 CC tetrapeptide associated with an oligopeptide which allows the
 CC introduction of the tetrapeptide into the target cell. The present
 CC sequence is one such peptide. The side chain linked to the lysine residue
 CC at position 1 corresponds to the specific tetrapeptide, while residues 1
 CC to 17 correspond to the oligopeptide. The specific tetrapeptide of the
 CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
 CC organisation in stress fibers. The present sequence may be used in the
 CC prevention and/or treatment of a disease related to alpha-SM actin
 CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
 CC fibrotic conditions. The present sequence may also be used to treat
 CC Dupuytren disease and lung fibrosis. The tetrapeptide of the present
 CC invention is described in AAB14263

XX Sequence 17 AA;

Query Match 72.8%; Score 75; DB 3; Length 17;
 Best Local Similarity 81.2%; Pred. No. 0.00046;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPKWK 17
 |||||
 Db 2 RQKIWFNRRKPKWK 17

RESULT 23

AAB13432
 ID AAB13432 standard; peptide; 18 AA.

XX

AC AAB13432;
 XX 23-NOV-2000 (first entry)
 XX
 XX Synthetic alpha smooth muscle actin inhibitor # 12.
 XX
 XX Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Cross_link 1
 FT /note= "Pentapeptide of Acetyl-EEEDC is linked via a
 FT disulphide bond to the thiol group of the side chain of
 FT the cysteine residue"
 XX
 XX WO200038733-A1.
 XX
 XX 06-JUL-2000.
 XX
 XX 15-DEC-1999; 99WO-EP009964.
 XX
 XX 24-DEC-1998; 98EP-00204396.
 XX
 XX (UNIO) UCB-BIOPRODUCTS SA.
 XX
 XX Gabbiani G, Scarso A;
 XX WPI; 2000-452308/39.
 XX
 XX A peptidic product for prevention and treatment of a disease related to
 PT alpha-SM actin expression comprises a tetrapeptide associated with a
 PT chemical entity that is able to introduce the tetrapeptide into the cell.
 XX
 XX Disclosure; Page; 31pp; English.
 XX
 XX The present invention relates to novel peptides comprising of a specific
 CC tetrapeptide associated with an oligopeptide which allows the
 CC introduction of the tetrapeptide into the target cell. The present
 CC sequence is one such peptide. The side chain linked to the lysine residue
 CC at position 1 corresponds to the specific tetrapeptide, while residues 1
 CC to 17 correspond to the oligopeptide. The specific tetrapeptide of the
 CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
 CC organisation in stress fibers. The present sequence may be used in the
 CC prevention and/or treatment of a disease related to alpha-SM actin
 CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
 CC Dupuytren disease and lung fibrosis. Note: the present sequence is not
 CC shown in the specification but is derived from information given by the
 CC inventors
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 72.8%; Score 75; DB 3; Length 18;
 Best Local Similarity 81.2%; Pred. No. 0.00049;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RQKIWFPPNRRKPKWK 17
 |||||
 Db 3 RQKIWFPPNRRKPKWK 18
 |||||
 RESULT 24
 AAB13428
 ID AAB13428 standard; peptide; 19 AA.
 XX
 XX AAB13428;
 AC
 XX 23-NOV-2000 (first entry)
 DT
 XX Synthetic alpha smooth muscle actin inhibitor # 8.

XX
 KW Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Biotinylated Gly"
 FT Modified-site 3
 FT /note= "Tetrapeptide of Acetyl-EEED linked on the epsilon
 FT -amino group of lysyl residue"
 XX
 XX WO200038733-A1.
 XX
 XX 06-JUL-2000.
 XX
 XX 15-DEC-1999; 99WO-EP009964.
 XX
 XX 24-DEC-1998; 98EP-00204396.
 XX
 XX (UNIO) UCB-BIOPRODUCTS SA.
 XX
 XX Gabbiani G, Scarso A;
 XX WPI; 2000-452308/39.
 XX
 XX A peptidic product for prevention and treatment of a disease related to
 PT alpha-SM actin expression comprises a tetrapeptide associated with a
 PT chemical entity that is able to introduce the tetrapeptide into the cell.
 XX
 XX Claim 7; Page 24; 31pp; English.
 XX
 XX The present invention relates to novel peptides comprising of a specific
 CC tetrapeptide associated with an oligopeptide which allows the
 CC introduction of the tetrapeptide into the target cell. The present
 CC sequence is one such peptide. The side chain linked to the lysine residue
 CC at position 3 corresponds to the specific tetrapeptide, while residues 3
 CC to 19 correspond to the oligopeptide. The specific tetrapeptide of the
 CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
 CC organisation in stress fibers. The present sequence may be used in the
 CC prevention and/or treatment of a disease related to alpha-SM actin
 CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
 CC fibrotic conditions. The present sequence may also be used to treat
 CC Dupuytren disease and lung fibrosis. The tetrapeptide of the present
 CC invention is described in AAB14263
 XX
 XX Sequence 19 AA;
 SQ
 Query Match 72.8%; Score 75; DB 3; Length 19;
 Best Local Similarity 81.2%; Pred. No. 0.00052;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RQKIWFPPNRRKPKWK 17
 |||||
 Db 4 RQKIWFPPNRRKPKWK 19
 |||||
 RESULT 25
 AAB13431
 ID AAB13431 standard; peptide; 20 AA.
 XX
 XX AAB13431;
 AC
 XX 23-NOV-2000 (first entry)
 DT
 XX Synthetic alpha smooth muscle actin inhibitor # 11.
 XX
 XX Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX

OS Synthetic.
XX
XX Key Location/Qualifiers
EH Modified-site 1
FT /note= "N-terminal acetyl"
FT
XX
XX WO200038733-A1.
XX
XX PD 06-JUL-2000.
XX
XX PF 15-DEC-1999; 99WO-EP009964.
XX
XX PR 24-DEC-1998; 98EP-00204396.
XX
XX (UNIO) UCB-BIOPRODUCTS SA.
XX
XX Gabbiani G, Scarso A;
XX
XX WPI; 2000-452308/39.
XX
XX A peptidic product for prevention and treatment of a disease related to
PT alpha-SM actin expression comprises a tetrapeptide associated with a
PT chemical entity that is able to introduce the tetrapeptide into the cell.
XX
XX
XX PS Disclosure; Page; 31pp; English.
XX
XX CC The present invention relates to novel peptides comprising of a specific
CC tetrapeptide associated with an oligopeptide which allows the
CC introduction of the tetrapeptide into the target cell. The present
CC sequence is one such peptide. Residues 1 to 4 of the present sequence
CC correspond to the specific tetrapeptide, while residues 5 to 20
CC correspond to the oligopeptide. The specific tetrapeptide of the present
CC sequence interferes with alpha smooth muscle (alpha-SM) actin
CC organisation in stress fibers. The present sequence may be used in the
CC prevention and/or treatment of a disease related to alpha-SM actin
CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
CC fibrotic conditions. The present sequence may also be used to treat
CC Dupuytren disease and lung fibrosis. Note: the present sequence is not
CC shown in the specification but is derived from information given by the
CC inventors
XX
XX SQ Sequence 20 AA;
Query Match 72.8%; Score 75; DB 3; Length 20;
Best Local Similarity 81.2%; Fred. NO. 0.00054;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 RQKIWFNRRKPKK 17
Db 5 RQKIWFNRRKPKK 20
Search completed: May 24, 2004, 17:22:20
Job time : 46.4865 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:18:06 ; Search time 9.87838 Seconds

(without alignments)
165.539 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IRQFKIWFPNRRKPKWK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 78.*

1: pirl.*

2: pirl2.*

3: pirl3.*

4: pirl4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	68	66.0	60	2	PC2399
2	68	66.0	60	2	PC2400
3	68	66.0	60	2	I51437
4	68	66.0	62	2	B32391
5	68	66.0	66	2	S15534
6	68	66.0	67	2	B27471
7	68	66.0	69	2	S13785
8	68	66.0	70	2	B37042
9	68	66.0	81	2	A25180
10	68	66.0	85	2	I65198
11	68	66.0	85	2	A25472
12	68	66.0	95	2	B32830
13	68	66.0	96	2	A52278
14	68	66.0	99	2	A41605
15	68	66.0	104	2	D43559
16	68	66.0	114	2	A43559
17	68	66.0	155	2	S25846
18	68	66.0	188	2	I50145
19	68	66.0	242	1	WJMSX3
20	68	66.0	243	1	WJMSX2
21	68	66.0	285	1	A43556
22	68	66.0	289	2	A43562
23	68	66.0	309	1	S09257
24	68	66.0	320	1	A39724
25	67	65.0	33	2	S57235
26	67	65.0	42	2	I65241
27	67	65.0	45	2	PC1216
28	67	65.0	48	2	I51439
29	67	65.0	66	2	S15536

30	67	65.0	66	2	S15538	homeotic protein H
31	67	65.0	71	2	JC1161	homeotic protein 3
32	67	65.0	71	2	A60084	homeotic protein H
33	67	65.0	74	2	D34510	homeotic protein H
34	67	65.0	75	2	I51341	homeo box protein
35	67	65.0	75	2	S58852	homeotic protein S
36	67	65.0	76	2	C43559	homeotic protein R
37	67	65.0	78	2	I51342	homeo box protein
38	67	65.0	81	2	S47605	homeotic protein H
39	67	65.0	81	2	B29585	homeotic protein H
40	67	65.0	82	2	S08302	homeotic protein H
41	67	65.0	83	2	S47603	homeotic protein H
42	67	65.0	83	2	S50066	homeotic protein H
43	67	65.0	86	2	A34510	homeotic protein H
44	67	65.0	86	2	JT0489	homeotic protein Z
45	67	65.0	86	2	S08303	homeotic protein H
46	67	65.0	87	2	S00589	homeotic protein H
47	67	65.0	88	2	A03317	homeotic protein M
48	67	65.0	96	2	S08639	homeotic protein z
49	67	65.0	96	2	A05266	homeotic protein H
50	67	65.0	97	2	C27176	homeotic protein H
51	67	65.0	97	2	A24779	homeotic protein m
52	67	65.0	103	2	A32167	homeotic protein H
53	67	65.0	105	2	S47602	homeotic protein H
54	67	65.0	105	2	A27471	homeotic protein R
55	67	65.0	106	2	S36448	homeotic protein s
56	67	65.0	107	2	B61045	homeobox protein t
57	67	65.0	113	2	T10775	homeotic protein H
58	67	65.0	118	2	A24777	homeotic protein H
59	67	65.0	118	2	JT0273	homeotic protein H
60	67	65.0	118	2	B24777	homeotic protein M
61	67	65.0	119	2	A03314	homeotic protein m
62	67	65.0	138	2	S20087	homeotic protein b
63	67	65.0	148	2	PC4071	homeobox A5 protei
64	67	65.0	153	1	WJHU3C	homeotic protein H
65	67	65.0	153	1	WJMSX6	homeotic protein H
66	67	65.0	158	2	A27348	homeotic protein H
67	67	65.0	209	2	A43553	homeotic protein H
68	67	65.0	217	1	WJHU2C	homeotic protein H
69	67	65.0	217	1	WJMSX2	homeotic protein H
70	67	65.0	220	2	A37371	homeotic protein H
71	67	65.0	220	2	S01063	Hox 2 protein type
72	67	65.0	222	2	S20029	homeotic protein H
73	67	65.0	224	2	S26400	homeotic protein H
74	67	65.0	224	2	A31324	homeotic protein H
75	67	65.0	227	2	S50067	homeotic protein H
76	67	65.0	228	2	S32563	homeotic protein H
77	67	65.0	229	2	A28329	homeotic protein H
78	67	65.0	230	2	S00592	homeotic protein H
79	67	65.0	232	1	A25108	homeotic protein H
80	67	65.0	232	2	S48125	homeotic protein H
81	67	65.0	233	2	I65137	hox1.3 protein - r
82	67	65.0	234	1	S02014	homeotic protein H
83	67	65.0	234	2	S00992	homeotic protein H
84	67	65.0	235	2	A56568	homeotic protein H
85	67	65.0	235	2	S72429	homeotic protein H
86	67	65.0	236	1	S09256	homeotic protein H
87	67	65.0	242	1	A39164	homeotic protein H
88	67	65.0	245	1	S10092	homeotic protein H
89	67	65.0	246	2	T46446	hypothetical prote
90	67	65.0	250	1	A36170	homeotic protein H
91	67	65.0	250	1	A31757	homeotic protein H
92	67	65.0	251	1	B60492	homeotic protein H
93	67	65.0	255	1	WJHU4B	homeotic protein H
94	67	65.0	264	1	WJHU3E	homeotic protein H
95	67	65.0	264	1	S35219	homeotic protein H
96	67	65.0	269	1	A43551	homeotic protein H
97	67	65.0	269	2	A45578	homeotic protein H
98	67	65.0	270	1	WJHU1C	homeotic protein H
99	67	65.0	270	1	WJMS13	homeotic protein H
100	67	65.0	275	1	WJZFX2	homeotic protein H

ALIGNMENTS

RESULT 1

PC2399
antennapedia-like homeotic protein AHox 2 - sea squirt (*Styela clava*) (fragment)
C/Species: *Styela clava*
C/Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jul-2000
C/Accession: PC2399
R/Ge, T.; Lee, H.; Tomlinson, C.R.
Gene 147; 219-222, 1994
A/Title: Identification of an antennapedia-like homeobox gene in the ascidians *Styela clava*
A/Reference number: PC2399; MUID:9501617; PMID:7926803
A/Accession: PC2399
A/Molecule type: DNA
A/Residues: 1-60 <GET>
A/Cross-references: GB:S73920; NID:G693714; PIDN:AAB33061.2; PID:G7387472
A/Note: The authors translated the codon ATT for residue 47 as Glu
C/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/2-58/Domain: homeobox homology <Hox>

Query Match	66.0%;	Score 68;	DB 2;	Length 60;
Best Local Similarity	75.0%;	Pred. No. 0.0027;		
Matches	12;	Conservative	0;	Mismatches 4;
			Indels	0;
			Gaps	0;

QY 2 ROPKIWFNPERRKPWK 17
 ||| ||||| |||||
Db 43 RQVKIWFQNRRMKWKK 58

RESULT 2

PC2400
antennapedia-like homeotic protein AHox 3 - sea squirt (*Styela plicata*) (fragment)
C/Species: *Styela plicata*
C/Date: 26-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Oct-1997
C/Accession: PC2400
R/Ge, T.; Lee, H.; Tomlinson, C.R.
Gene 147; 219-222, 1994
A/Title: Identification of an antennapedia-like homeobox gene in the ascidians *Styela plicata*
A/Reference number: PC2399; MUID:9501617; PMID:7926803
A/Accession: PC2400
A/Molecule type: DNA
A/Residues: 1-60 <GET>
A/Note: The authors translated the codon ATA for residue 47 as Glu
C/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F/2-58/Domain: homeobox homology <Hox>

Query Match	66.0%;	Score 68;	DB 2;	Length 60;
Best Local Similarity	75.0%;	Pred. No. 0.0027;		
Matches	12;	Conservative	0;	Mismatches 4;
			Indels	0;
			Gaps	0;

QY 2 ROPKIWFNPERRKPWK 17
 ||| ||||| |||||
Db 43 RQVKIWFQNRRMKWKK 58

RESULT 3

I51437
homeotic protein Hox 2.4 - African clawed frog (fragment)
C/Species: *Xenopus laevis* (African clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C/Accession: I51437
R/Bittner, D.
Dev. Dyn. 196, 11-24, 1993
A/Title: Characterization of the Xenopus Hox 2.4 gene and identification of control elements
A/Reference number: I51437; MUID:93326800; PMID:7916675
A/Accession: I51437
A/Status: preliminary; translated from GB/EMBL/DDBB
A/Molecule type: DNA
A/Residues: 1-60 <BIT>
A/Cross-references: GB:M83947; NID:q214252; PIDN:AAAA9748.1; PID:g214253

RESULT 6

B27471 homeotic protein R2 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 17-Nov-2000
 C:Accession: B27471; I65243
 R:Palzon, M.; Sanderson, N.; Chung, S.Y.
 Gene 54, 23-32, 1987

A:Title: Cloning and expression of rat homeo-box-containing sequences.

A:Reference number: A91576; MUID:87277429; PMID:2886401

A:Accession: B27471

A:Molecule type: DNA

A:Residues: 1-67 <FAL>

R:Sakoyama, Y.; Mizuta, I.; Ogasawara, N.; Yoshikawa, H.

Biochem. Genet. 32, 351-360, 1994

A:Title: Cloning of rat homeobox genes.

A:Reference number: I52340; MUID:95217128; PMID:7702549

A:Accession: I65243

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-42 <RES>

A:Cross-references: GB:S76296; NID:g913079

C:Genetics:

A:Gene: Hox-A; Hox-1

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 66.0%; Score 68; DB 2; Length 67;

Best Local Similarity 75.0%; Pred. No. 0.003;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17

Db 25 RQVKWFPQNRMRKWK 40

RESULT 7

S13785 homeotic protein m31 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 17-Oct-1997
 R:Breier, G.; Bucan, M.; Francke, U.; Colberg-Poley, A.M.; Gruss, P.
 EMBO J. 5, 2209-2215, 1986

A:Title: Sequential expression of murine homeo box genes during F9 EC cell differentiation

A:Reference number: S13785; MUID:87053860; PMID:2877873

A:Accession: S13785

A:Molecule type: DNA

A:Residues: 1-69 <BRE>

C:Genetics:

A:Map position: 15

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;2-58/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 2; Length 69;

Best Local Similarity 75.0%; Pred. No. 0.0031;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17

Db 43 RQVKWFPQNRMRKWK 58

RESULT 8

B37042 homeotic protein Hox B8 - human (fragment)
 N:Alternate names: homeotic protein Hox 2.4; homeotic protein Hox 2D
 C:Species: Homo sapiens (man)
 C>Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 24-Sep-1999
 C:Accession: B37042; S15533

R:Giampaolo, A.; Acampora, D.; Zappavigna, V.; Pannese, M.; D'Esposito, M.; Care, A.; Fa
 Differentiation 40, 191-197, 1989
 A:Title: Differential expression of human HOX-2 genes along the anterior-posterior axis
 A:Reference number: A37042; MUID:89378558; PMID:2570724

A:Accession: B37042

A:Molecule type: DNA

A:Residues: 1-70 <GIA>

A:Cross-references: GB:X16173; NID:g32372; PIDN:CAA34295.1; PID:g930067

R:Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Stc:

Genome 31, 745-756, 1989

A:Title: Organization of human class I homeobox genes.

A:Reference number: S15036; MUID:90215256; PMID:2576652

A:Accession: S15533

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 5-70 <BON>

C:Genetics:

A:Gene: GDB:HOXB8

A:Cross-references: GDB:120661; OMIM:142963

A:Map position: 17q21.3-17q21.3

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;6-62/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 2; Length 70;

Best Local Similarity 75.0%; Pred. No. 0.0032;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17

Db 47 RQVKWFPQNRMRKWK 62

RESULT 9

A25180 homeotic protein Hox 3 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 31-Dec-1993 #sequence_revision 03-Feb-1994 #text_change 24-Sep-1999
 R:Awgulewitsch, A.; Utset, M.F.; Hart, C.P.; McGinnis, W.; Ruddle, F.H.
 Nature 320, 328-335, 1986
 A:Title: Spatial restriction in expression of a mouse homeo box locus within the centra.
 A:Reference number: A25180; MUID:86175026; PMID:3007994

A:Accession: B25180

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <AWG>

A:Cross-references: GB:X03659; NID:g51401; PIDN:CAA27294.1; PID:g929686

A>Note: the authors translated the codon CAA for residue 47 as Glu

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;16-72/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 2; Length 81;

Best Local Similarity 75.0%; Pred. No. 0.0036;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWK 17

Db 57 RQVKWFPQNRMRKWK 72

RESULT 10

I65198 homeotic protein Hox A4 - rat (fragment)
 N:Alternate names: homeotic protein Hox 1.4
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 02-Aug-1996 #sequence_revision 23-May-1997 #text_change 20-Aug-1999
 C:Accession: I65198
 R:Gorski, D.H.; LePage, D.F.; Walsh, K.
 BioTechniques 16, 856-858, 1994
 A:Title: Cloning and sequence analysis of homeobox transcription factor cDNAs with an inc
 A:Reference number: I52196; MUID:94347374; PMID:7915120

C;Genetics:
A;Gene: GDB:HOXD8
A;Cross-references: GDB:120680; OMIM:142985
A;Map position: 2q31-2q31
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;6-62/Domain: homeobox homology <Hox>

Query Match 66.0%; Score 68; DB 2; Length 95;
Best Local Similarity 75.0%; Pred. No. 0.0042; 4; Gaps 0;
Matches 12; Conservative 0; Mismatches 4; Indels 0;

QY 2 RQPKIWFPNRRKPKWK 17
||| ||| ||| ||| |||
DB 47 RQVKIWFQNRMRKWK 62

RESULT 13
A55278
homeotic protein HOX3A - cat (fragment)
C;Species: Felis silvestris catus (domestic cat)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Oct-1997
C;Accession: A55278
R;Masuda, R.; Yuhki, N.; O'Brien, S.J.
Genomics 11, 1007-1013, 1991
A;Title: Molecular cloning, chromosomal assignment, and nucleotide sequence of the felid
A;Reference number: A55278; MUID:92147095; PMID:1686012
A;Accession: A55278
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-96 <MAS>
A;Note: sequence extracted from NCBI backbone (NCBIN:80214, NCBI:80215)
C;Genetics:
A;Gene: HOX3A
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;4-60/Domain: homeobox homology <Hox>

Query Match 66.0%; Score 68; DB 2; Length 96;
Best Local Similarity 75.0%; Pred. No. 0.0043; 4; Gaps 0;
Matches 12; Conservative 0; Mismatches 4; Indels 0;

QY 2 RQPKIWFPNRRKPKWK 17
||| ||| ||| ||| |||
DB 45 RQVKIWFQNRMRKWK 60

RESULT 14
A41605
homeotic protein Hox 4.3 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1992 #sequence_revision 01-Dec-1995 #text_change 24-Sep-1999
C;Accession: S16177; C42694; A41605; A38810; S15521
R;Sadoul, R.; Featherstone, M.S.
Biochim. Biophys. Acta 1089, 259-261, 1991
A;Title: Sequence analysis of the homeobox-containing exon of the murine Hox-4.3 homeoge
A;Reference number: S16177; MUID:91274361; PMID:1675873
A;Accession: S16177
A;Molecule type: DNA
A;Residues: 1-99 <BIO>
A;Cross-references: EMBL:X56561; NID:G51418; PIDN:CAA39911.1; PID:e30697; PID:gl333935
R;Nazarali, A.; Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 89, 2883-2887, 1992
A;Title: Hox-1.11 and Hox-4.9 homeobox genes.
A;Reference number: A42694; MUID:92212934; PMID:1348361
A;Accession: C42694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2-70 <NAZ>
A;Cross-references: GB:M87803; NID:gl93953; PIDN:AAA37852.1; PID:gl93954
A;Note: sequence extracted from NCBI backbone (NCBIN:92289, NCBI:92290)
R;Singh, G.; Kaur, S.; Stock, J.L.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Potter
Proc. Natl. Acad. Sci. U.S.A. 89, 10706-10710, 1991

2 RQPKIWFPNRRKPWKK 17

RESULT 19
WJMSX3
homeotic protein Hox C8 - mouse
NAlternate names: homeotic protein Hox 3.1
C.Species: Mus musculus (house mouse)
C.Date: 20-Jun-1991 #sequence revision 30-1

C;Accession: B36023; S00548; A29980
 R;Angewitsch, A.; Bieberich, C.; Bogarad, L.; Shashikant, C.; Ruddle, F.H.
 Proc Natl Acad Sci U.S.A. 87, 6428-6432, 1990
 A;Title: Structural analysis of the Hox-3.1 transcription unit and the Hox-3.2-Hox-3.1
 A;Reference number: A36023; MUID:90349629; PMID:1696731
 A;Accession: B36023
 A;Molecule type: DNA
 A;Residues: 1-242 <AWG>
 A;Cross-references: GB:M35603; NID:g193977; PIDN:AAA37857.1; PID:g193979
 R;Breier, G.; Dressler, G.R.; Gruss, P.
 EMBO J. 7, 1329-1336, 1988
 A;Title: Primary structure and developmental expression pattern of Hox 3.1, a member of
 A;Reference number: S00548; MUID:88312579; PMID:2900757
 A;Accession: S00548
 A;Molecule type: mRNA
 A;Residues: 1-242 <BRE>
 A;Cross-references: EMBL:X07646; NID:g51402; PIDN:CAA30486.1; PID:g51403
 R;Le Mouellic, H.; Condamine, H.; Brulet, P.
 Genes Dev. 2, 125-135, 1988
 A;Title: Pattern of transcription of the homeo gene Hox-3.1 in the mouse embryo.
 A;Reference number: A29980; MUID:88185818; PMID:2895723
 A;Accession: A29980
 A;Molecule type: mRNA
 A;Residues: 1-242 <LEN>
 A;Cross-references: EMBL:X07439; NID:g51406; PIDN:CAA30319.1; PID:g51407
 C;Genetics:
 A;Gene: Hox 3.1
 A;Map position: 15
 A;Introns: 146/1
 C;Superfamily: homeotic protein Hox A7; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;150-206/Domain: homeobox homology <Hox>

Query Match 66.0%; Score 68; DB 1; Length 242;
 Best Local Similarity 75.0%; Pred. No. 0.01;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKIWFNRRKPKWK 17
 || ||||| |||||
 Db 191 RQVKIWFQNRMMKWK 206

RESULT 20
 WUMS24
 N;Alternate names: homeotic protein Hox B8 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 22-Jun-1999
 R;Kongsuwan, K.; Allen, J.; Adams, J.M.
 Nucleic Acids Res. 17, 1881-1892, 1989
 A;Title: Expression of Hox-2.4 homeobox gene directed by proviral insertion in a myeloid
 A;Reference number: S03712; MUID:89183599; PMID:2564662
 A;Accession: S03712
 A;Molecule type: mRNA
 A;Residues: 1-243 <KON>
 A;Cross-references: EMBL:X13721
 R;Blatt, C.
 submitted to the EMBL Data Library, July 1989
 A;Reference number: S16691
 A;Accession: S16691
 A;Molecule type: DNA
 A;Residues: 1-88, 'D', '90-243 <BLAI>
 A;Cross-references: EMBL:X13961; NID:g51393; PIDN:CAA32141.1; PID:g295919
 R;Blatt, C.; Aberdam, D.; Schwartz, R.; Sachs, L.
 EMBO J. 7, 4283-4290, 1988
 A;Title: DNA rearrangement of a homeobox gene in myeloid leukaemic cells.
 A;Reference number: S02015; MUID:89210815; PMID:2907477
 A;Accession: S02015
 A;Molecule type: DNA
 A;Residues: 1-29, 'HDLPATHRGVRSQORHVPFANPGVLPRAIVAVHSSLPABFVRRGCN', 82-88, 'D', '90-114, 'R
 A;Cross-references: EMBL:X13961
 A;Note: this sequence has been revised in reference S16691

R;Hart, C.P.; Fainsod, A.; Ruddle, F.H.
 Genomics 1, 182-195, 1987
 A;Title: Sequence analysis of the murine Hox-2.2, -2.3, and -2.4 homeo boxes: evolutiona
 A;Reference number: A27176; MUID:88085193; PMID:2891608
 A;Accession: A27176
 A;Molecule type: DNA
 A;Residues: 143-243 <HAR>
 A;Cross-references: EMBL:M18399
 C;Genetics:
 A;Gene: Hox-2.4
 A;Map position: 11
 A;Introns: 142/1
 C;Superfamily: homeotic protein Hox A7; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;147-203/Domain: homeobox homology <Hox>

Query Match 66.0%; Score 68; DB 1; Length 243;
 Best Local Similarity 75.0%; Pred. No. 0.01;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKIWFNRRKPKWK 17
 || ||||| |||||
 Db 188 RQVKIWFQNRMMKWK 203

RESULT 21
 A43556
 homeotic protein Hox A4 - mouse
 N;Alternate names: homeotic protein 1.3; homeotic protein HBT-1; homeotic protein Hox 1.
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Dec-1992 #sequence revision 11-Dec-1992 #text_change 22-Jun-1999
 C;Accession: A43556; B43556; A25311; A25107; I49753; S26147
 R;Galliot, B.; Dolle, P.; Vigneron, M.; Featherstone, M.S.; Baron, A.; Duboule, D.
 Development 107, 343-359, 1989
 A;Title: The mouse Hox-1.4 gene: primary structure, evidence for promoter activity and e
 A;Reference number: A43556; MUID:90214520; PMID:2576648
 A;Accession: A43556
 A;Molecule type: DNA
 A;Residues: 1-285 <GAL>
 A;Cross-references: EMBL:X66861
 A;Accession: B43556
 A;Molecule type: mRNA
 A;Residues: 1-285 <GA2>
 R;Rubin, M.R.; Toth, L.E.; Patel, M.D.; D'Eustachio, P.; Nguyen-Huu, M.C.
 Science 233, 663-667, 1986
 A;Title: A mouse homeo box gene is expressed in spermatocytes and embryos.
 A;Reference number: A25311; MUID:86261825; PMID:3726554
 A;Accession: A25311
 A;Molecule type: mRNA
 A;Residues: 180, 'E', 182-209, 211-268, 'H', 270-271 <RUB>
 A;Cross-references: GB:M13813
 R;Wolgemuth, D.J.; Engelmyer, E.; Duggal, R.N.; Gizang-Ginsberg, E.; Mutter, G.L.; Ponze
 EMBO J. 5, 1229-1235, 1986
 A;Title: Isolation of a mouse cDNA coding for a developmentally regulated, testis-specif
 A;Reference number: A25107; MUID:86274625; PMID:2426103
 A;Accession: A25107
 A;Molecule type: mRNA
 A;Residues: 198-268, 'H', 270-271 <WOL>
 A;Cross-references: GB:M27432
 R;Duggal, R.N.; Zakeri, Z.F.; Ponsetto, C.; Wolgemuth, D.J.
 Ann. N. Y. Acad. Sci. 513, 112-127, 1987
 A;Title: Differential expression of the c-abl proto-oncogene and the homeo box-containin
 A;Reference number: I49753; MUID:88181884; PMID:2895600
 A;Accession: I49753
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 180-268, 'H', 270-285 <RES>
 A;Cross-references: GB:M27432; NID:g193957; PIDN:AAA16440.1; PID:g193958
 C;Genetics:
 A;Gene: Hoxa-4; Hox 1.4; MH-3
 A;Map position: 6
 A;Introns: 171/1
 C;Function:

A;Description: control of embryonic development by tissue- and stage-specific regulation
 C;Superfamily: homeotic protein Hox D4; homeobox homology
 C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
 F;181-237/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 1; Length 285;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
 ||||| ||||| |||||

Db 222 RQVKIWFQNNRRMKWK 237
 ||||| ||||| |||||

RESULT 22

A43562
 homeotic protein Hox D8 - mouse
 N;Alternate names: homeotic protein Hox 4.3
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-May-1997
 C;Accession: A43562
 R;Izpisua-Belmonte, J.C.; Dolle, P.; Renucci, A.; Zappavigna, V.; Falkenstein, H.; Dubou
 Development 110, 733-745, 1990
 A;Title: Primary structure and embryonic expression pattern of the mouse Hox-4.3 homeob
 A;Reference number: A43562; MUID:91209232; PMID:1982431
 A;Accession: A43562
 A;Status: Preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-289 <I2P>
 C;Superfamily: homeotic protein Hox A7; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;196-252/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 2; Length 289;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
 ||||| ||||| |||||

Db 237 RQVKIWFQNNRRMKWK 252
 ||||| ||||| |||||

RESULT 23

S09257
 homeotic protein Hox A4 - chicken
 N;Alternate names: homeotic protein Chox1-4
 C;Species: Gallus gallus (chicken)
 C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000
 C;Accession: S09257; S10883
 R;Sasaki, H.; Yokoyama, E.; Kuroiwa, A.
 Nucleic Acids Res. 18, 1739-1747, 1990
 A;Title: Specific DNA binding of the two chicken deformed family homeodomain proteins, C
 A;Reference number: S09256; MUID:90245562; PMID:1970866
 A;Accession: S09257
 A;Molecule type: mRNA
 A;Residues: 1-309 <SAS>
 A;Cross-references: EMBL:X52670; NID:G632118; PIDN:CAA36896.1; PID:G63219
 R;Scotting, P.J.; Hewitt, M.; Keynes, R.J.
 Nucleic Acids Res. 18, 3999, 1990
 A;Title: Isolation and analysis of chick homeobox cDNA clones.
 A;Reference number: S10883; MUID:90326535; PMID:1973835
 A;Accession: S10883
 A;Molecule type: mRNA
 A;Residues: 207-273 <SCO>
 A;Cross-references: EMBL:X52747; NID:G63223; PIDN:CAB57949.1; PID:G6018426
 C;Function:

A;Description: control of embryonic development by tissue- and stage-specific regulation
 C;Superfamily: homeotic protein Hox D4; homeobox homology
 C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
 F;210-266/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 1; Length 309;
 Best Local Similarity 75.0%; Pred. No. 0.013;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RQPKIWFNNRRKPKWK 17
 ||||| ||||| |||||
 Db 251 RQVKIWFQNNRRMKWK 266
 ||||| ||||| |||||

RESULT 24

A39724
 homeotic protein Hox A4 - human
 N;Alternate names: homeotic protein Hox 1.4; homeotic protein Hox 1D
 C;Species: Homo sapiens (man)
 C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 26-Feb-1999
 C;Accession: A39724; A60492; S15544; A32468
 R;Buettnner, R.; Yim, S.O.; Hong, Y.S.; Boncinelli, E.; Tainsky, M.A.
 Mol. Cell. Biol. 11, 3573-3583, 1991
 A;Title: Alteration of homeobox Gene expression by N-ras transformation of PA-1 human te
 A;Reference number: A39724; MUID:91260707; PMID:1675427
 A;Accession: A39724
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-320 <BUE>
 A;Cross-references: GB:M39724
 R;Peverali, F.A.; D'Esposito, M.; Acampora, D.; Bunone, G.; Negri, M.; Faiella, A.; Stori
 Differentiation 45, 61-89, 1990
 A;Title: Expression of HOX homeogenes in human neuroblastoma cell culture lines.
 A;Reference number: A60492; MUID:91153613; PMID:1981366
 A;Accession: A60492
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-107, 'A', 109-320 <PEV>
 R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Stoi
 Genome 31, 745-756, 1989
 A;Title: Organization of human class I homeobox genes.
 A;Reference number: S15036; MUID:90215256; PMID:2576652
 A;Accession: S15544
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 215-280 <BON>
 R;Ferguson-Smith, A.C.; Fienberg, A.; Ruddle, F.H.
 Genomics 5, 250-258, 1989
 A;Title: Isolation, chromosomal localization, and nucleotide sequence of the human HOX 1.
 A;Reference number: A32468; MUID:90007544; PMID:2571574
 A;Accession: A32468
 A;Molecule type: DNA
 A;Residues: 215-275 <FER>
 A;Cross-references: EMBL:M28199
 C;Comment: This homeotic protein is expressed in embryonic nervous system.

C;Genetics:
 A;Gene: GDB:HOXA4
 A;Cross-references: GDB:120650; OMIM:142953
 A;Map position: 7p15.3-7p15.3
 C;Function:
 A;Description: control of embryonic development by tissue- and stage-specific regulation
 C;Superfamily: homeotic protein Hox D4; homeobox homology
 C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
 F;216-272/Domain: homeobox homology <HOX>

Query Match 66.0%; Score 68; DB 1; Length 320;
 Best Local Similarity 75.0%; Pred. No. 0.013;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
 ||||| ||||| |||||

Db 257 RQVKIWFQNNRRMKWK 272
 ||||| ||||| |||||

RESULT 25

S57235
 antennapedia protein (clone p1105) - fruit fly (Drosophila pseudoobscura) (fragment)
 C;Species: Drosophila pseudoobscura
 C;Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
 C;Accession: S57235

R;Randazzo, F.M.; Seeger, M.A.; Huss, C.A.; Sweeney, M.A.; Cecil, J.K.; Kaufman, T.C.
 Genetics 133, 319-330, 1993
 A;Title: Structural changes in the antennapedia complex of *Drosophila pseudoobscura*.
 A;Reference number: S57224
 A;Accession: S57235
 A;Molecule type: DNA
 A;Residues: 1-33 <RAN>
 A;Cross-references: EMBL:X77711
 C;Genetics:
 A;Gene: FlyBase:Antp
 A;Cross-references: FlyBase:FBgn0012693
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;1-2/Domain: homeobox homology (fragment) <HOX>

Query Match 65.0%; Score 67; DB 2; Length 33;
 Best Local Similarity 75.0%; Pred. No. 0.0022; 4; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 4

QY 2 RPKIWFPPNRRKPKKK 17
 |||||
 Db 7 RQIKIWFQNRRMKWKK 22

Search completed: May 24, 2004, 17:26:11
 Job time : 10.8784 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:17:30 ; Search time 6.2027 Seconds
(without alignments)
142.711 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IRQPKIWFPPNRKPKWK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	66.0	67	1 HXA4 RAT	P09635 rattus norv
2	68	66.0	92	1 HXB8 RAT	P18863 rattus norv
3	68	66.0	108	1 HXB8 RAT	P18866 rattus norv
4	68	66.0	188	1 HXD8 CHICK	P23459 gallus gall
5	68	66.0	240	1 HXD8 HETFR	Q91a12 heterodontu
6	68	66.0	242	1 HXC8 HUMAN	P31273 homo sapien
7	68	66.0	242	1 HXC8 MOUSE	P09025 mus musculu
8	68	66.0	243	1 HXB8 HUMAN	P17481 homo sapien
9	68	66.0	243	1 HXB8 MOUSE	P09632 mus musculu
10	68	66.0	247	1 HXA4 HETFR	Q91a22 heterodontu
11	68	66.0	248	1 HXA4 MORSA	Q9pvd2 morone saxa
12	68	66.0	285	1 HXA4 MOUSE	P06798 mus musculu
13	68	66.0	289	1 HXD8 MOUSE	P23463 mus musculu
14	68	66.0	290	1 HXD8 HUMAN	P13378 homo sapien
15	68	66.0	309	1 HXA4 CHICK	P17277 gallus gall
16	68	66.0	320	1 HXA4 HUMAN	Q00056 homo sapien
17	67	65.0	48	1 HXB6 XENLA	P31256 xenopus lae
18	67	65.0	49	1 HXA5 SHEEP	Q28599 ovis aries
19	67	65.0	71	1 HXA7 SHEEP	Q28500 ovis aries
20	67	65.0	71	1 HXC5 NOTVI	P31262 notophthalm
21	67	65.0	74	1 HMO9 APIME	P15860 apis mellif
22	67	65.0	75	1 HMO9 SALSA	P09636 salmo salar
23	67	65.0	76	1 HXC4 RAT	P18865 rattus norv
24	67	65.0	78	1 HXA5 SALSA	P09637 salmo salar
25	67	65.0	80	1 HXA4 LNSA	P81192 lineus sang
26	67	65.0	81	1 HXS5 BRARE	P09013 brachydanio
27	67	65.0	82	1 HXB5 CHICK	P14838 gallus gall
28	67	65.0	84	1 HXB6 CHICK	P14839 gallus gall
29	67	65.0	86	1 SCR APIME	P15859 apis mellif
30	67	65.0	87	1 HXC5 XENLA	P09020 xenopus lae
31	67	65.0	93	1 HXB8 PIG	P09078 sus scrofa
32	67	65.0	96	1 HXC6 BRARE	P15862 brachydanio
33	67	65.0	105	1 HXA7 RAT	P09634 rattus norv

34	67	65.0	105	1 HXB4 BRARE	P22574 brachydanio
35	67	65.0	112	1 HXB7 RAT	P18864 rattus norv
36	67	65.0	148	1 HXA5 AMBME	P50208 ambystoma m
37	67	65.0	153	1 HXC6 SHEEP	P49925 ovis aries
38	67	65.0	208	1 HXA7 HETFR	Q91a25 heterodontu
39	67	65.0	209	1 HXA7 XENLA	P09071 xenopus lae
40	67	65.0	217	1 HXB7 BOVIN	Q91t89 bos taurus
41	67	65.0	217	1 HXB7 HUMAN	P09629 homo sapien
42	67	65.0	217	1 HXB7 MOUSE	P09024 mus musculu
43	67	65.0	220	1 HBA7 XENLA	Q91771 xenopus lae
44	67	65.0	220	1 HBA7 XENLA	Q91771 xenopus lae
45	67	65.0	222	1 HXC5 HUMAN	Q04444 homo sapien
46	67	65.0	222	1 HXC5 MOUSE	P17043 mus musculu
47	67	65.0	224	1 HXB6 HUMAN	P17509 homo sapien
48	67	65.0	224	1 HXB6 MOUSE	P09023 mus musculu
49	67	65.0	225	1 HXA7 MORSA	Q9pvd4 morone saxa
50	67	65.0	228	1 HXB6 BRARE	P15861 brachydanio
51	67	65.0	229	1 HXA6 HETFR	Q91a24 heterodontu
52	67	65.0	229	1 HXA6 MOUSE	P02830 mus musculu
53	67	65.0	230	1 HXA7 HUMAN	P31268 homo sapien
54	67	65.0	230	1 HXB5 XENLA	P09019 xenopus lae
55	67	65.0	232	1 HXA6 MOUSE	P09092 mus musculu
56	67	65.0	232	1 HXB4 XENLA	P09070 xenopus lae
57	67	65.0	232	1 HXC5 BRARE	P09074 brachydanio
58	67	65.0	233	1 HXA5 RAT	P52949 rattus norv
59	67	65.0	233	1 HXA6 HUMAN	P31267 homo sapien
60	67	65.0	234	1 HXC6 NOTVI	P14858 notophthalm
61	67	65.0	234	1 HXC6 XENLA	P02832 xenopus lae
62	67	65.0	235	1 HXC6 HUMAN	P09630 homo sapien
63	67	65.0	235	1 HXC6 MOUSE	P10629 mus musculu
64	67	65.0	235	1 HXD4 CHICK	P17278 gallus gall
65	67	65.0	236	1 HXD4 BRARE	O57374 brachydanio
66	67	65.0	242	1 HXA7 COTUA	P24061 coturnix co
67	67	65.0	245	1 HXB4 CHICK	P14840 gallus gall
68	67	65.0	250	1 HXB4 MOUSE	P10284 mus musculu
69	67	65.0	250	1 HXD4 MOUSE	P10628 mus musculu
70	67	65.0	251	1 HXB4 FUGRU	O13074 fugu rubrip
71	67	65.0	251	1 HXB4 HUMAN	P17483 homo sapien
72	67	65.0	252	1 HXD5 HETFR	Q91a11 heterodontu
73	67	65.0	255	1 HXD4 HUMAN	P09016 homo sapien
74	67	65.0	261	1 HXC4 ORYLA	Q9pvs4 oryzias lat
75	67	65.0	264	1 HXC4 HUMAN	P09017 homo sapien
76	67	65.0	264	1 HXC4 MOUSE	O08624 mus musculu
77	67	65.0	269	1 HXB5 HUMAN	P09067 homo sapien
78	67	65.0	269	1 HXB5 MOUSE	P09079 mus musculu
79	67	65.0	270	1 HXA5 HUMAN	P09021 mus musculu
80	67	65.0	270	1 HXA5 MOUSE	P09014 heterodontu
81	67	65.0	275	1 HXA5 HETFR	Q91a23 brachydanio
82	67	65.0	275	1 HXB5 BRARE	Q9pvd3 morone saxa
83	67	65.0	281	1 HXA5 MORSA	P02833 drosophila
84	67	65.0	378	1 HMAN DROSU	Q24645 drosophila
85	67	65.0	394	1 SCR DROME	P09077 drosophila
86	67	65.0	415	1 HMD3 DROME	P07548 drosophila
87	67	65.0	590	1 HMD3 TRIGR	P10178 tripneustes
88	64	62.1	108	1 HMB3 TRIGR	P32442 mus musculu
89	64	62.1	253	1 MOX1 HUMAN	P50221 homo sapien
90	64	62.1	254	1 MOX1 MOUSE	P39021 xenopus lae
91	64	62.1	298	1 MOX2 XENLA	P50222 homo sapien
92	64	62.1	303	1 MOX2 HUMAN	P32443 mus musculu
93	64	62.1	303	1 MOX2 MOUSE	P39020 rattus norv
94	64	62.1	303	1 MOX2 CAEL	P56407 caenorhabdi
95	63	61.2	60	1 HMO9 BRARE	P15858 apis mellif
96	63	61.2	74	1 HMB5 APIME	Q28598 ovis aries
97	63	61.2	87	1 HMO1 SHEEP	P53547 caenorhabdi
98	63	61.2	132	1 HMO1 CAEL	P19601 gallus gall
99	63	61.2	232	1 SAX1 CHICK	P42580 mus musculu
100	63	61.2	305	1 SAX1 MOUSE	

ALIGNMENTS

RESULT 1

```

HX44_RAT
ID HX44_RAT STANDARD; PRT; 67 AA.
AC P09635;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (R2) (Fragment).
GN HOXA4 OR HOK-A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=8727429; PubMed=2886401;
RA Falzon M., Sanderson N., Chung S.Y.;
RT "Cloning and expression of rat homeo-box-containing sequences.";
RT Gene 54:23-32(1987).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RT Biochem. Genet. 32:351-360(1994).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16808; -; NOT_ANNOTATED_CDS.
CC F01; B27471; B27471.
CC HSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD00010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1 HOMEBOX.
CC DNA_BIND <1 42
CC NON_TER 67 67
CC SEQUENCE 67 AA; 8346 MW; 5BFDFAE93193780 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 67;
Best Local Similarity 75.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIFPNRRKPKWK 17
DB 25 RQVKIFQNRMRKWK 40

RESULT 2
HX44_RAT
ID HX44_RAT STANDARD; PRT; 92 AA.
AC P1863;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (R2) (Fragment).
GN HOXA4 OR HOK-A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=8727429; PubMed=2886401;
RA Falzon M., Sanderson N., Chung S.Y.;
RT "Cloning and expression of rat homeo-box-containing sequences.";
RT Gene 54:23-32(1987).
RN [2]
RP SEQUENCE OF 1-42 FROM N.A.
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RT Biochem. Genet. 32:351-360(1994).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16808; -; NOT_ANNOTATED_CDS.
CC F01; B27471; B27471.
CC HSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD00010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1 HOMEBOX.
CC DNA_BIND <1 42
CC NON_TER 67 67
CC SEQUENCE 67 AA; 8346 MW; 5BFDFAE93193780 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 67;
Best Local Similarity 75.0%; Pred. No. 0.00033;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIFPNRRKPKWK 17
DB 25 RQVKIFQNRMRKWK 40

RESULT 3
HX44_RAT
ID HX44_RAT STANDARD; PRT; 108 AA.
AC P1886;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C8 (R4) (Fragment).
GN HOXC8 OR HOKC-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

DT 01-NOV-1990 (Rel. 16, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B8 (R1A) (Fragment).
GN HOXB8 OR HOKB-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=89231502; PubMed=2907739;
RA Falzon M., Chung S.Y.;
RT "The expression of rat homeobox-containing genes is developmentally
RT regulated and tissue specific.";
RL Development 103:601-610(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M37565; AAA41341.1; ALT_SEQ.
CC F01; A43559; A43559.
CC HSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambdarepressor.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSOR.
CC ProDom; PD00010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1 HOMEBOX.
CC DNA_BIND 11 70
CC NON_TER 92 92
CC SEQUENCE 92 AA; 11358 MW; 3D8E3EA42622B4CA CRC64;

Query Match 66.0%; Score 68; DB 1; Length 92;
Best Local Similarity 75.0%; Pred. No. 0.00046;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIFPNRRKPKWK 17
DB 53 RQVKIFQNRMRKWK 68

RESULT 3
HX44_RAT
ID HX44_RAT STANDARD; PRT; 108 AA.
AC P1886;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C8 (R4) (Fragment).
GN HOXC8 OR HOKC-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RL Biochem. Genet. 32:351-360(1994).
[2]
RN SEQUENCE OF 5-108 FROM N.A.
RP STRAIN=Sprague-Dawley;
RC MEDLINE=89231502; PubMed=2907739;
RA Falzon M., Chung S.Y.;
RX "The expression of rat homeobox-containing genes is developmentally
RT regulated and tissue specific.";
RL Development 103:601-610(1988).
CC -i- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -i- SUBCELLULAR LOCATION: Nuclear.
CC -i- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -i- SIMILARITY: Belongs to the Antp homeobox family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; S76301; ; NOT ANNOTATED_CDS.
DR EMBL; M37568; AAA41344.1; -.
DR PIR; D43559; D43559.
DR HSP; P02833; 9ANT.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESS.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR PROSITE; PS00027; HOMEBOX; 1; 1.
DR PROSITE; PS00071; HOMEBOX; 2; 1.
DR PROSITE; PS00071; HOMEBOX; 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON_TER 1
FT DNA_BIND 15 74 HOMEBOX.
FT SEQUENCE 108 AA; 13115 MW; DBF8C0AD6A0B08E1 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 108;
Best Local Similarity 75.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
DB 57 RQVKIWFQNRKPKWK 72

RESULT 4
HXD8_CHICK
ID HXD8_CHICK STANDARD; PRT; 188 AA.
AC P23459;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D8 (Chox-M).
GN HOXD8 OR CHOX-M.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

```

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91238215; PubMed=1674560;
RA Crompton M.R., McGregor A.D., Goodwin G.H.;
RT "cDNA cloning of a homeobox-containing gene expressed in avian
RT myeloblastic virus-transformed chicken monoblastic leukaemia cells.";
RL Leukemia 5:357-360(1991).
CC -i- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -i- SUBCELLULAR LOCATION: Nuclear.
CC -i- SIMILARITY: Belongs to the Antp homeobox family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X57158; CAA40445.1; -.
DR PIR; I50145; I50145.
DR HSP; P02833; 9ANT.
DR TRANSPAC; T01754; -.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESS.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX; 1; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00071; HOMEBOX; 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 15 25 POLY-ALA.
FT SITE 83 88 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 95 154 HOMEBOX.
FT SEQUENCE 188 AA; 21729 MW; D4560E8807FE29FE CRC64;

Query Match 66.0%; Score 68; DB 1; Length 188;
Best Local Similarity 75.0%; Pred. No. 0.00096;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
DB 137 RQVKIWFQNRKPKWK 152

RESULT 5
HXD8_HETFR
ID HXD8_HETFR STANDARD; PRT; 240 AA.
AC Q91A12;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D8.
GN HOXD8.
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OX NCBI_TaxID=7792;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20144096; PubMed=10677514;

```


RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,
RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;
RT "Hox cluster genomics in the horn shark, *Heterodontus francisci*.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF224263; AAF44632.1; -.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T04482; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW transcription regulation.
FT SITE 135 140 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 147 206 HOMEBOX.
FT SEQUENCE 240 AA; 28127 MW; 4B9CC0022CE58105 CRC64;
SQ
Query Match 56.0%; Score 68; DB 1; Length 240;
Best Local Similarity 75.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 RQPKIWFPPNRKPKWK 17
DB 189 RQVKIWFQNRKPKWK 204

RESULT 6
ID HX8 HUMAN STANDARD; PRT; 242 AA.
AC P12273; O15221; O15362;
CT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein Hox-C8 (Hox-3A).
GN HOC8 OR HOC3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RA Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.;
RT "A complete mutation analysis panel of human HOX genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RP TTSUB=Eye;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Small D.E.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 64-114 AND 147-195 FROM N.A.
RX MEDLINE=98019011; PubMed=9357979;
RA Flagiello D., Gibaud A., Dutrillaux B., Poupon M.F., Malfoy B.;
RT "Distinct patterns of all-trans retinoic acid dependent expression of
RT HOXB and HOXC homeogenes in human embryonal and small-cell lung
RT carcinoma cell lines.";
RL FEBS Lett. 415:263-267(1997).
RN [4]
RP SEQUENCE OF 149-214 FROM N.A.
RX MEDLINE=90215256; PubMed=2576652;
RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RL Genome 31:745-756(1989).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBUNIT: INTERACTS WITH SMAD1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY014300; AAG42146.1; -.
CC EMBL; AY014299; AAG42146.1; JOINED.
CC EMBL; BC053898; AAG53898.1; -.
CC EMBL; X99680; CAA67996.1; -.
CC EMBL; X99681; CAA67997.1; -.
CC FIR; S15534; S15534.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T03326; -.
CC Genew; HGNC:5129; HOXC8.
CC MIM; 142970; -.
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003700; F:transcription factor activity; NAS.
CC GO; GO:0007275; P:development; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambrapressr.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.

```

DR PROSITE; PS50071; HOMEBOX 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 138 143 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 149 208 HOMEBOX.
FT CONFLICT 165 165 E -> Q (IN REF. 4).
SQ SEQUENCE 242 AA; 27754 MW; 30C1FD06228833FC CRC64;

Query Match 66.0%; Score 68; DB 1; Length 242;
Best Local Similarity 75.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
DB 191 RQVKIWFNNRRKPKWK 206

RESULT 7
HXCB MOUSE
ID HXCB MOUSE STANDARD; PRT; 242 AA.
AC P09025;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-CB (Hox-3.1) (M31).
GN HOXC8 OR HOXC-8 OR Hox-3.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SE SEQUENCE FROM N.A.
RP MEDLINE=88185618; PubMed=2895723;
RA le Mouellie H., Condamine H., Brulet P.;
RT "Pattern of transcription of the homeo gene Hox-3.1 in the mouse
RT embryo.";
RL Genes Dev. 2:125-135(1988).
RN [2]
SE SEQUENCE FROM N.A.
RP MEDLINE=88312579; PubMed=2900757;
RA Breier G., Dressler G.R., Gruss P.;
RT "Primary structure and developmental expression pattern of Hox 3.1, a
RT member of the murine Hox 3 homeobox gene cluster.";
RL EMBO J. 7:1329-1336(1988).
RN [3]
SE SEQUENCE FROM N.A.
RP STRAIN=CD-1; TISSUE=Connective tissue;
RX MEDLINE=90349629; PubMed=1696731;
RA Angulewitsch A., Bieberich C., Bogard L., Shashikant C., Ruddle F.H.;
RT "Structural analysis of the Hox-3.1 transcription unit and the
RT Hox-3-2-Hox-3.1 intergenic region.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6428-6432(1990).
RN [4]
SE SEQUENCE OF 147-216 FROM N.A.
RP MEDLINE=86175026; PubMed=3007994;
RA Angulewitsch A., Utset M.F., Hart C.P., McGinnis W., Ruddle F.H.;
RT "Spatial restriction in expression of a mouse homoeo box locus within
RT the central nervous system.";
RL Nature 320:328-335(1986).
RN [5]
SE SEQUENCE OF 149-215 FROM N.A.
RP MEDLINE=87053860; PubMed=2877873;
RA Breier G., Bucan M., Francke U., Colberg-Poley A.M., Gruss P.;
RT "Sequential expression of murine homeo box genes during F9 EC cell
RT differentiation.";
RL EMBO J. 5:2209-2215(1986).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: Initially found in all tissues of the
CC posterior region in 8.5 and 9.5 dpc. Embryos, it eventually become
CC specifically located in neural tissue.

```

```

CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; X03659; CAA27294.1; ALT_SEQ.
EMBL; X07646; CAA30486.1; -.
EMBL; X07439; CAA30319.1; -.
EMBL; M35603; AAA37857.1; -.
PIR; B36023; WJMSX3.
PIR; S13785; S13785.
HGSP; P02833; 9ANT.
DR TRANSFAC; T01749; -.
DR MGD; MGI:96198; Hoxc8.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepreser.
DR Pfam; PF00045; homeobox_1.
DR PRINTS; PR00025; ANTENNAPEIDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 138 143 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 149 208 HOMEBOX.
FT CONFLICT 151 151 S -> T (IN REF. 4).
FT CONFLICT 181 181 E -> Q (IN REF. 4).
SQ SEQUENCE 242 AA; 27740 MW; 4A2A53BDA4CFA878 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 242;
Best Local Similarity 75.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
DB 191 RQVKIWFNNRRKPKWK 206

RESULT 8
HXCB HUMAN
ID HXCB HUMAN STANDARD; PRT; 243 AA.
AC P17481; Q9H112;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B8 (Hox-2D) (Hox-2.4).
GN HOXB8 OR HOXD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SE SEQUENCE FROM N.A.
RP Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.;
RT "A complete mutation analysis panel of human HOX genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SE SEQUENCE OF 142-211 FROM N.A.
RP TISSUE=Placenta; PubMed=2570724;
RX MEDLINE=89378558; PubMed=2570724;
RA Giampaolo A., Acampora D., Zappavigna V., Pannese M.,
RA D'Esposito M., Care A., Faiella A., Stornaiuolo A., Russo G.,
RA Simone A., Boncinelli E., Peschle C.;

```

RT "Differential expression of human HOX-2 genes along the anterior-posterior axis in embryonic central nervous system.";
 RL Differentiation 40:191-197(1989).
 RN [3]
 RP SEQUENCE OF 146-211 FROM N.A.
 RX MEDLINE=90215256; PubMed=2576652;
 RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
 RA Gaudino G., Stornaiuolo A., Carlieri M., Faiella A., Simeone A.;
 RT "Organization of human class I homeobox genes.";
 RL Genome 31:745-756(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC a specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DEVELOPMENTAL STAGE: THE HOX-2 GENES ARE EXPRESSED IN
 CC WHOLE EMBRYOS AND FETUSES AT 5-9 WEEKS FROM CONCEPTION.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY014294; AAG42143.1; -;
 DR EMBL; AY014293; AAG42143.1; JOINED.
 DR EMBL; X16173; CAA34295.1; -;
 DR PIR; B37042; B37042.
 DR HSP; P02833; 9ANT.
 DR TRANSFAC; T03324; -;
 DR Genew; HGNC:5119; HOXB8.
 DR MIM; 142963; -;
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH lambdarepressor.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00031; HTHREPRESS.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT SITE 134 139 ANT-P-TYPE HEXAPEPTIDE.
 FT DNA BIND 146 205 HOMEBOX.
 FT SEQUENCE 243 AA; 27573 MW; 36982182725F121F CRC64;
 Query Match 66.0%; Score 68; DB 1; Length 243;
 Best Local Similarity 75.0%; Pred. No. 0.0013;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RQKWFPPNRRKPKWK 17
 Db 188 RQKWFPPNRRKPKWK 203
 RESULT 9
 HXB8_MOUSE
 ID HXB8_MOUSE STANDARD; PRT; 243 AA.
 AC P09632;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-B8 (Hox-2.4).

GN HOXB8 OR HOXB-8 OR HOX-2.4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BALEB/C;
 RC MEDLINE=99183599; PubMed=2564662;
 RX Kongsuan K., Allen J., Adams J.M.;
 RA "Expression of Hox-2.4 homeobox gene directed by proviral insertion
 RT in a myeloid leukemia.";
 RL Nucleic Acids Res. 17:1881-1892(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BALEB/C;
 RC MEDLINE=89210815; PubMed=2907477;
 RX Blatt C., Aberdam D., Schwartz R., Sachs L.;
 RA "DNA rearrangement of a homeobox gene in myeloid leukaemic cells.";
 RT EMBO J. 7:4283-4290(1988).
 RN [3]
 RP REVISIONS.
 RA Blatt C.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJF databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220677; PubMed=1673811;
 RX Ben-David L., Aberdam D., Sachs L., Blatt C.;
 RA "A deletion and a rearrangement distinguish between the
 RT intracisternal A- particle of Hox-2.4 and that of interleukin-3 in
 RT the same leukemic cells.";
 RL Virology 182:382-387(1991).
 RN [5]
 RP SEQUENCE OF 144-243 FROM N.A.
 RX MEDLINE=88085193; PubMed=2891608;
 RA Hart C.P., Fainsod A., Ruddle F.H.;
 RT "Sequence analysis of the murine Hox-2.2, -2.3, and -2.4 homeo boxes:
 RT evolutionary and structural comparisons.";
 RL Genomics 1:182-195(1987).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X13961; CAA32141.1; -;
 DR EMBL; X13721; CAA32000.1; -;
 DR EMBL; M18399; AAA88246.1; ALT_SEQ.
 DR EMBL; X54077; CAA38014.1; -;
 DR PIR; S03712; WTMS24.
 DR HSP; P02833; 9ANT.
 DR TRANSFAC; T01737; -;
 DR MGI; MGI:96189; Hoxb8.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH lambdarepressor.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESS.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT SITE 134 139 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 146 205 HOMEBOX.
 SQ SEQUENCE 243 AA; 27618 MW; 368964C77884D21F CRC64;

Query Match 66.0%; Score 68; DB 1; Length 243;
 Best Local Similarity 75.0%; Pred. No. 0.0013;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
 DB 188 RQVKIWFQNRKMKWK 203

RESULT 10

HXA4_HETFR
 ID HXA4_HETFR STANDARD; PRT; 247 AA.
 AC Q91A22;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-A4.
 GN HOXA4.
 OS Heterodontus francisci (Horn shark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Heterodontoidea; Heterodontiformes;
 OC Heterodontidae; Heterodontus.
 OX NCBI_TaxID=7792;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2014096; PubMed=10677514;
 RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,
 RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;
 RT "Hox cluster genomics in the horn shark, Heterodontus francisci.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AF224262; AAF44642.1; --
 DR HSSP; P02833; 9ANT.
 DR TRANSFAC; T04475; --
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEIDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT SITE 128 133 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 149 208 HOMEBOX.
 SQ SEQUENCE 247 AA; 28408 MW; BE518038CA73D2E6 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 247;
 Best Local Similarity 75.0%; Pred. No. 0.0013;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RQPKIWFNRRKPKWK 17
 DB 191 RQVKIWFQNRKMKWK 206

RESULT 11

HXA4_MORSA
 ID HXA4_MORSA STANDARD; PRT; 248 AA.
 AC Q9PWD2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-A4.
 GN HOXA4.
 OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99259633; PubMed=10327649;
 RA Snell E.A., Scemama J.L., Stellwag E.J.;
 RT "Genomic organization of the Hoxa4-Hoxa10 region from Morone
 saxatilis: implications for Hox gene evolution among vertebrates.";
 RL J. Exp. Zool. 285:41-49(1999).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AF089743; AAD46399.1; --
 DR HSSP; P02833; 9ANT.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEIDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT SITE 128 133 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 149 208 HOMEBOX.
 SQ SEQUENCE 248 AA; 27729 MW; DAGE182A92655F13 CRC64;

Query Match 66.0%; Score 68; DB 1; Length 248;
 Best Local Similarity 75.0%; Pred. No. 0.0013;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
 DB 191 RQVKIWFQNRKMKWK 206

RESULT 12

HXA4 MOUSE
ID HXA4 MOUSE STANDARD; PRT; 285 AA.
AC P06798; Q61684; Q64388;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (Hox-1.4) (MH-3).
GN HXA4 OR HOXA-4 OR Hox-1.4.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-170 FROM N.A.
RX MEDLINE=90214520; PubMed=2576648;
RA Galliot B., Dolle P., Vigneron M., Featherstone M.S., Baron A.,
Duboule D.;
RT "The mouse Hox-1.4 gene: primary structure, evidence for promoter
activity and expression during development.";
RL Development 107:343-359(1989).
RN [2]
RP SEQUENCE OF 171-285 FROM N.A.
RC STRAIN=ICR; TISSUE=Spinal cord, and Testis;
RX MEDLINE=922190549; PubMed=1686835;
RA Rubin M.R., Nguyen-Huu M.C.;
RT "Murine embryonic spinal cord and adult testis Hox-1.4 cDNAs are
RT identical 3' to the homeo box.";
RL DNA Seq. 1:329-334(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94227844; PubMed=7909702;
RA Viviano C.M., Galliot B., Wolgemuth D.J.;
RT "Multiple levels of regulation exist for expression of the Hoxa-4
RT (Hox-1.4) gene in the mouse testis.";
RL Cell. Mol. Biol. Res. 39:483-495(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=94043512; PubMed=7901228;
RA Wu K., Wolgemuth D.J.;
RT "Protein product of the somatic-type transcript of the Hoxa-4 (Hox-
RT 1.4) gene binds to homeobox consensus binding sites in its promoter
RT and intron.";
RL J. Cell. Biochem. 52:449-462(1993).
RN [5]
RP SEQUENCE OF 180-273 FROM N.A.
RX MEDLINE=86261825; PubMed=3726554;
RA Rubin M.R., Toth L.E., Patel M.D., D'Eustachio P., Nguyen-Huu M.C.;
RT "A mouse homeo box gene is expressed in spermatocytes and embryos.";
RL Science 233:663-667(1986).
RN [6]
RP SEQUENCE OF 198-271 FROM N.A.
RX MEDLINE=86274625; PubMed=2426103;
RA Wolgemuth D.J., Engelmyer E., Duggal R.N., Gizang-Ginsberg E.,
Matter G.L., Ponsetto C., Viviano C., Zakeri Z.F.;
RT "Isolation of a mouse cDNA coding for a developmentally regulated,
RT testis-specific transcript containing homeo box homology.";
RL EMBO J. 5:1229-1235(1986).
RN [7]
RP SEQUENCE OF 180-285 FROM N.A.
RX MEDLINE=88181884; PubMed=2895600;
RA Duggal R.N., Zakeri Z.F., Ponsetto C., Wolgemuth D.J.;
RT "Differential expression of the c-abl proto-oncogene and the homeo
RT box-containing gene Hox 1.4 during mouse spermatogenesis.";
RL Ann. N.Y. Acad. Sci. 513:112-127(1987).
RN [8]
RP RESULTS OF OVEREXPRESSION
RX MEDLINE=89127494; PubMed=2563568;
RA Wolgemuth D.J., Behringer R.R., Mostoller M.P., Brinster R.L.,
Palmiter R.D.;
RT "Transgenic mice overexpressing the mouse homeobox-containing gene
RT Hox-1.4 exhibit abnormal gut development.";
RL Nature 337:464-467(1989).
RN [9]
RP FUNCTION: Sequence-specific transcription factor which is part of

a developmental regulatory system that provides cells with
specific positional identities on the anterior-posterior axis.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- DEVELOPMENTAL STAGE: DURING DEVELOPMENT OF THE PREPUBERAL TESTIS
HIGH LEVELS OF Hox-1.4 TRANSCRIPTS WERE FOUND AT DAYS 17, 24 AND
30. THE FIRST DAY OF Hox-1.4 EXPRESSION WAS DAY 14. THE ACTIVATION
OF THE Hox-1.4 GENE IN MALE GERM CELLS SEEMS TO OCCUR AT THE
PACHYTENE STAGE OF MEIOTIC PROPHASE AND ITS LEVEL OF EXPRESSION IS
STAGE-SPECIFIC DURING EMBRYOGENESIS.
-!- DISEASE: Overexpression results in abnormal gut development
(megacolon).
-!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X66861; CAA47330.1; -;
DR EMBL; X17346; CAA35228.1; -;
DR EMBL; X13538; CAA31889.1; -;
DR EMBL; S70444; AAB30705.2; -;
DR EMBL; S67058; AAB28662.2; -;
DR EMBL; M13813; AAA37831.1; -;
DR EMBL; M27432; AAA16440.1; -;
DR PIR; A43556; A43556.
DR HSSP; P02833; 9ANT.
DR TRANSFAC; T01701; -;
DR MGD; MGI:96176; Hoxa4.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation.
KW SITE 159 164 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 180 239 HOMEBOX.
FT CONFLICT 97 97 G -> A (IN REF. 1).
FT CONFLICT 181 181 K -> E (IN REF. 2; CAA35228 AND 5).
FT CONFLICT 210 210 MISSING (IN REF. 5).
FT CONFLICT 269 269 H -> P (IN REF. 3 AND 4).
FT CONFLICT 272 273 HP -> PE (IN REF. 5).
FT CONFLICT 279 285 TRIPSSI -> HPFPFPYNELEWISFSLCALLISAPVLIC
SSLSKPEPEKQPTCKWK (IN REF. 2).
FT SQ SEQUENCE 285 AA; 30467 MW; 475D48D6C6302A26 CRC64;
Query Match 66.0%; Score 68; DB 1; Length 285;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 RQPKWFPNRRKPKWK 17
DB 222 RQVKWFPNRRKPKWK 237

RESULT 13
HXD8 MOUSE
ID HXD8 MOUSE STANDARD; PRT; 289 AA.
AC P23453;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D8 (Hox-4.3) (Hox-5.4).

GN HOXD8 OR HOXD-8 OR HOX-4.3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91209232; PubMed=1982431;
 RA Ispisua-Belmonte J.-C., Dolle P., Renucci A., Zappavigna V.,
 RA Falkenstein H., Duboule D.;
 RT "Primary structure and embryonic expression pattern of the mouse
 RT Hox-4.3 homeobox gene.";
 RL Development 110:733-745(1990).
 RN [2]
 RP SEQUENCE OF 191-289 FROM N.A.
 RX MEDLINE=91274361; PubMed=1675873;
 RA Sadoul R., Featherstone M.;
 RA "Sequence analysis of the homeobox-containing exon of the murine
 RT Hox-4.3 homeogene.";
 RL Biochim. Biophys. Acta 1089:259-261(1991).
 RN [3]
 RP SEQUENCE OF 195-254 FROM N.A.
 RX MEDLINE=92073356; PubMed=1683707;
 RA Singh G., Kaur S., Stock J.L., Jenkins N.A., Gilbert D.J.,
 RA Copeland N.G., Potter S.S.;
 RT "Identification of 10 murine homeobox genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710(1991).
 RN [4]
 RP SEQUENCE OF 192-260 FROM N.A.
 RX MEDLINE=92212934; PubMed=1348361;
 RA Nazarali A., Kim Y., Nirenberg M.;
 RA "Hox-1.11 and Hox-4.9 homeobox genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2883-2887(1992).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.

FT CONFLICT 265 266 EA -> DG (IN REF. 1).
 FT CONFLICT 275 A -> V (IN REF. 2).
 SQ SEQUENCE 289 AA; 31410 MW; 5783099FB9B2BDFE CRC64;
 Query Match 56.0%; Score 68; DB 1; Length 289;
 Best Local Similarity 75.0%; Pred. No. 0.0015;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ROPKIWFNNRRKPWK 17
 DB 237 RQVKIWFNNRRKPWK 252
 RESULT 14
 HXD8 HUMAN STANDARD; PRT; 290 AA.
 AC P13378;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-D8 (Hox-4E) (Hox-5.4).
 GN HOXD8 OR HOX4E.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Birren B., Linton L., Nuebaum C., Lander E.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.;
 RT "A complete mutation analysis panel of human HOX genes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 193-287 FROM N.A.
 RX MEDLINE=89306602; PubMed=2568311;
 RA Oliver G., Sidell N., Fiske N., Heinzmann C., Mohandas T.,
 RA Sparkes R.S., de Robertis E.M.;
 RT "Complementary homeo protein gradients in developing limb buds.";
 RL Genes Dev. 3:641-650(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AC009336; -; NOT ANNOTATED_CDS.
 EMBL; AY014304; AAG42152.1; -.
 EMBL; AY014303; AAG42152.1; JOINED.
 EMBL; X15507; CAA33529.1; -.
 FIR; B32830; B32830.
 HSSP; P02833; 9ANT.
 TRANSFAC; T03332; -.
 Genew; HGNC:5139; HOXD8.
 MIM; 142985; -.
 GO; GO:0005634; C:nucleus; NAS.
 GO; GO:0003700; F:transcription of anterior/posterior axis, e...; NAS.
 GO; GO:0008595; P:determination of transcription, DNA-dependent; NAS.
 GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 InterPro; IPR001827; Antennapedia.
 InterPro; IPR001356; Homeobox.
 Pfam; PF00046; Homeobox; 1.
 PRINTS; PR00025; ANTENNAPEDIA.
 T -> S (IN REF. 1).
 T -> S (IN REF. 1).

```

DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 15 23 POLY-ALA.
FT DOMAIN 45 50 POLY-ALA.
FT DOMAIN 109 123 POLY-PRO.
FT DNA BIND 197 256 HOMEBOX.
FT CONFLICT 287 287 G -> A (IN REF. 3).
SQ SEQUENCE 290 AA; 31910 MW; 75FF95A73E2AA85F CRC64;

Query Match 66.0%; Score 68; DB 1; Length 290;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
DB 239 RQVKIWFQNRMRKWK 254

RESULT 15
HX44 CHICK STANDARD; PRT; 309 AA.
AC P17277;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (Chox-1.4).
GN HOXA4 OR CHOX-1.4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI TaxID=9031;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=90245562; PubMed=1970866;
RA Sasaki H., Yokoyama B., Kuroiwa A.;
RT "Specific DNA binding of the two chicken Deformed family homeodomain
RT proteins, Chox-1.4 and Chox-a.";
RL Nucleic Acids Res. 18:1739-1747(1990).
RN [2]
RP SEQUENCE OF 207-273 FROM N.A.
RC STRAIN=Comet Hubbard hybrid.
RX MEDLINE=90326535; PubMed=1973835;
RA Scotting P.J., Hewitt M., Keynes R.J.;
RT "Isolation and analysis of chick homeobox cDNA clones.";
RL Nucleic Acids Res. 18:3999-3999(1990).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- FUNCTION: BINDS TO SITES IN THE 5'-FLANKING SEQUENCE OF ITS CODING
CC REGION WITH VARIOUS AFFINITIES. THE CONSENSUS SEQUENCES OF THE
CC HIGH AND LOW AFFINITY BINDING SITES ARE TAATGA(C/G) AND CTAATTTT.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: THE PROLINE STRETCH WORKS AS A PART OF THE
CC TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X52670; CAA36896.1; -.

```

RA Ferguson-Smith A.C., Fienberg A., Ruddle F.H.;
 RT "Isolation, chromosomal localization, and nucleotide sequence of the
 RT human HOX 1.4 homeobox.";
 RL Genomics 5:250-258(1989);
 CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
 CC BINDS TO SITES IN THE 5'-FLANKING SEQUENCE OF ITS CODING REGION
 CC WITH VARIOUS AFFINITIES. THE CONSENSUS SEQUENCES OF THE HIGH AND
 CC LOW AFFINITY BINDING SITES ARE 5'-TAATGA[CG]-3' AND
 CC 5'-CTAATTT-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Embryonic nervous system.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M74297; AAA58664.1; -;
 DR EMBL; M28199; AAA53290.1; -;
 DR EMBL; AC004080; -; NOT ANNOTATED_CDS.
 DR PIR; A39724; A39724.
 DR HSP; P02833; 9ANT.
 DR TRANSFAC; T01703; -;
 DR Genew; HGNC:5105; HOXA4.
 DR GIM; 142953; -;
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEIDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT DOMAIN 15 192
 FT SITE 194 199
 FT DNA_BIND 215 274
 FT CONFLICT 70 70
 FT P -> T (IN REF. 3).
 FT A -> P (IN REF. 1).
 FT CONFLICT 108 108
 FT P -> L (IN REF. 3).
 FT CONFLICT 140 140
 FT A -> S (IN REF. 5).
 FT CONFLICT 205 205
 FT CONFLICT 275 276
 FT KL -> NC (IN REF. 5).
 SQ SEQUENCE 320 AA; 34479 MW; ASE5C8108F0646D CRC64;
 Query Match 66.0%; Score 68; DB 1; Length 320;
 Best Local Similarity 75.0%; Pred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 RQKIFPNRRKPKWK 17
 Db 257 RQKIFQNRMRKWK 272
 RESULT 17
 HXB6_XENLA
 ID HXB6_XENLA STANDARD; PRT; 48 AA.
 AC P31256;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-B6 (XlHox-2.2) (fragment).

GN HOXB6 OR XLHOX-2.2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93043517; PubMed=1384809;
 RA Leroy P., de Robertis E.M.;
 RT "Effects of lithium chloride and retinoic acid on the expression of
 RT genes from the Xenopus laevis Hox 2 complex.";
 RL Dev. Dyn. 194:21-32(1992).
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M91587; AAA49750.1; -;
 DR PIR; I51439; I51439.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; PARTIAL.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT NON_TER 1 1
 FT DNA_BIND <1 29
 FT SEQUENCE 48 AA; 5716 MW; BC39E36822EDDD2A CRC64;
 Query Match 65.0%; Score 67; DB 1; Length 48;
 Best Local Similarity 75.0%; Pred. No. 0.00033;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 RQKIFPNRRKPKWK 17
 Db 12 RQKIFQNRMRKWK 27
 RESULT 18
 HXA5_SHEEP
 ID HXA5_SHEEP STANDARD; PRT; 49 AA.
 AC Q28599;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-A5 (Fragment).
 GN HOXA5 OR HOXA-5.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Roche P.J.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with

CC specific positional identities on the anterior-posterior axis.
 CC Also binds to its own promoter. Binds specifically to the motif:
 CC 5'-CYNATTA[GG]Y-3'.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U61978; AAB04754.1; -.
 CC HSP; P02833; IHOM.
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX 1; 1.
 CC PROSITE; PS00071; HOMEBOX 2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC Transcription regulation.
 CC KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC FT NON_TER 1 1
 CC DNA_BIND <1 49
 CC NON_TER 49 49
 CC SEQUENCE 49 AA; 6331 MW; 1EE702315E7C099B CRC64;
 CC
 CC Query Match 65.0%; Score 67; DB 1; Length 49;
 CC Best Local Similarity 75.0%; Pred. No. 0.00034;
 CC Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC Qy 2 RQPKIWFNRRKPKWK 17
 CC ||||| |||||
 CC 32 RQIKWQNRKMKWK 47
 CC
 CC RESULT 19
 CC HXA7 SHEEP
 CC ID HXA7 SHEEP STANDARD; PRT; 71 AA.
 CC AC Q28600;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Homeobox protein Hox-A7 (Fragment).
 CC GN HOXA7 OR HOXA-7.
 CC OS Ovis aries (Sheep).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Pecora; Bovidae;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OC Bovidae; Caprinae; Ovis.
 CC OX NCBI_TaxID=9940;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Roche P.J.;
 CC RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U61979; AAB04755.1; -.
 CC HSP; P02833; 9ANT.

DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH lambrepressor.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 KW FT NON_TER 1 1
 FT DNA_BIND 4 63
 FT NON_TER 71 71
 FT SEQUENCE 71 AA; 8888 MW; 931049FAC1BAACB7 CRC64;
 SQ
 Query Match 65.0%; Score 67; DB 1; Length 71;
 Best Local Similarity 75.0%; Pred. No. 0.0005;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 RQPKIWFNRRKPKWK 17
 ||||| |||||
 46 RQIKWQNRKMKWK 61
 Db
 RESULT 20
 HXC5 NOTVI
 ID HXC5 NOTVI STANDARD; PRT; 71 AA.
 AC P31262;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-C5 (NvHox-3.4) (Fragment).
 OS Notochthalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
 OC Notochthalmus.
 OC NCBI_TaxID=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92290273; PubMed=1351019;
 RA Belleville S., Beauchemin M., Tremblay M., Noiseux N., Savard P.;
 RT "Homeobox-containing genes in the newt are organized in clusters
 RT similar to other vertebrates.";
 RL Gene 114:179-186(1992).
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M84001; AAB49397.1; ALT_INIT.
 CC FIR; JCI161; JCI161.
 CC HSP; P02833; ISAN.
 CC InterPro; IPR001827; Antennapedia.
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX 1; 1.
 CC PROSITE; PS00032; ANTENNAPEdia; PARTIAL.
 CC PROSITE; PS00071; HOMEBOX 2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.

```

FT  NON TER      1      1
FT  DNA BIND     4      63
FT  NON TER      71     71
SQ  SEQUENCE     71 AA; 8979 MW; 07999FDE89995B42 CRC64;

Query Match
Best Local Similarity 65.0%; Score 67; DB 1; Length 71;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2 ROPKIWFNRRKPKWK 17
DB  46 RQIKWIFQNRRMKWK 61

RESULT 21
HM90 APIME
ID  HM90 APIME STANDARD; PRT; 74 AA.
AC  P15860;
DT  01-APR-1990 (Rel. 14, Created)
DT  01-APR-1990 (Rel. 14, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Homeobox protein H90 (Fragment).
OS  Apis mellifera (Honeybee).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC  Apidae; Apis.
OX  NCBI_TaxID=7460;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90093384; PubMed=2574865;
RA  Walldorf U., Fleig R., Gehring W.J.;
RT  "Comparison of homeobox-containing genes of the honeybee and
RL  Drosophila.";
CC  Proc. Natl. Acad. Sci. U.S.A. 86:9971-9975(1989).
CC  -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC  -!- SIMILARITY: Belongs to the Antp homeobox family.
CC  -!- SIMILARITY: Contains 1 homeobox domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M29493; AAA27728.1; -.
DR  PIR; D34510; D34510.
DR  HSP; P02833; IHOM.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; Homeobox; 1.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT  NON TER      1      67
FT  DNA BIND     8      67
FT  NON TER      74     74
SQ  SEQUENCE     74 AA; 9263 MW; 5FC8FB4F723D3837 CRC64;

Query Match
Best Local Similarity 65.0%; Score 67; DB 1; Length 74;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2 ROPKIWFNRRKPKWK 17
DB  50 RQIKWIFQNRRMKWK 65

RESULT 22
HMSA_SALSA
ID  HMSA_SALSA STANDARD; PRT; 75 AA.
AC  P09636;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-MAR-1989 (Rel. 10, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Homeobox protein SI2-A (Fragment).
OS  Salmo salar (Atlantic salmon).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX  NCBI_TaxID=8030;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88226009; PubMed=2897318;
RA  Fjose A., Molven A., Eiken H.G.;
RT  "Molecular cloning and characterization of homeo-box-containing genes
RL  from Atlantic salmon.";
CC  Gene 62:141-152(1988).
CC  -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -!- SIMILARITY: Belongs to the Antp homeobox family.
CC  -!- SIMILARITY: Contains 1 homeobox domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M18903; AAA49559.1; -.
DR  PIR; I51341; I51341.
DR  HSP; P02833; 9ANT.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT  NON TER      1      60
FT  DNA BIND     1      60
FT  NON TER      75     75
SQ  SEQUENCE     75 AA; 9330 MW; FC02C3672F35475D CRC64;

Query Match
Best Local Similarity 65.0%; Score 67; DB 1; Length 75;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  2 ROPKIWFNRRKPKWK 17
DB  43 RQIKWIFQNRRMKWK 58

RESULT 23
HXC4_RAT
ID  HXC4_RAT STANDARD; PRT; 76 AA.
AC  P18865;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-NOV-1990 (Rel. 16, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Homeobox protein Hox-C4 (R3) (Fragment).
GN  HOXC4 OR HOXC-4.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=Sprague-Dawley;
RA  MEDLINE=89231502; PubMed=2907739;
RA  Falzon M., Chung S.Y.;

```

```

RT "The expression of rat homeobox-containing genes is developmentally
RL regulated and tissue specific.";
RC Development 103:601-610(1998).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M37567; AAA41343.1; -.
CC PIR; C43559; C43559.
CC HSSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC DNA_BIND 11 70 HOMEBOX.
CC SEQUENCE 76 AA; 9293 MW; 5235F665C0672385 CRC64;
CC
CC Query Match 65.0%; Score 67; DB 1; Length 76;
CC Best Local Similarity 75.0%; Pred. No. 0.00054;
CC Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 2 RQPKIWFNNRRKPKWK 17
CC ||||| |||||
CC Db 53 RQIKIWFQNNRRMKWK 68
CC
CC RESULT 24
CC HXA5_SALSA
CC ID HXA5_SALSA STANDARD; PRT; 78 AA.
CC AC P09637;
CC DT 01-MAR-1989 (Rel. 10, Created)
CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Homeobox protein Hox-A5 (S12-B) (Fragment).
CC GN HXA5.
CC OS Salmo salar (Atlantic salmon).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
CC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
CC NCBI_TaxID=8030;
CC [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=88226009; PubMed=2897318;
CC RA Fjose A., Molven A., Eiken H.G.;
CC RT "Molecular cloning and characterization of homeo-box-containing genes
CC from Atlantic salmon.";
CC Gene 62:141-152(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.

```

```

CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18904; AAA49560.1; -.
CC PIR; I51342; I51342.
CC HSSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC DNA_BIND 1 60 HOMEBOX.
CC SEQUENCE 78 AA; 9489 MW; 828DEBDDF78AC820 CRC64;
CC
CC Query Match 65.0%; Score 67; DB 1; Length 78;
CC Best Local Similarity 75.0%; Pred. No. 0.00055;
CC Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 2 RQPKIWFNNRRKPKWK 17
CC ||||| |||||
CC Db 43 RQIKIWFQNNRRMKWK 58
CC
CC RESULT 25
CC HXA4_LINSA
CC ID HXA4_LINSA STANDARD; PRT; 80 AA.
CC AC P81192;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Homeobox protein Hox-A4 (LsHox 4) (Fragment).
CC GN HXA4.
CC OS Lineus sanguineus (Ribbon worm).
CC OC Eukaryota; Metazoa; Nemertea; Anopla; Heteronemertea; Lineidae;
CC Lineus.
CC NCBI_TaxID=48190;
CC [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=98169491; PubMed=9501210;
CC RA Kmita-Cunisse M., Loeuf F., Bierre J., Gehring W.J.;
CC RT "Homeobox genes in the ribbonworm Lineus sanguineus: evolutionary
CC implications.";
CC Proc. Natl. Acad. Sci. U.S.A. 95:3030-3035(1998).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC HSSP; P02833; 9ANT.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC InterPro; IPR000047; HTH lambdarepressr.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.

```

DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS00071; HOMEOBOX_2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR HOMEOBOX; DNA-binding; Developmental protein; Nuclear protein;
KW transcription regulation.
FT NON_TER 1
FT DNA_BIND 21 80 HOMEOBOX.
FT NON_TER 80
SQ SEQUENCE 80 AA; 9860 MW; F2CE1B01CB8042F1 CRC64;

Search completed: May 24, 2004, 17:25:15
Job time : 6.2027 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:14:23 ; Search time 29.4054 Seconds
(without alignments)
182.409 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IRQKINFPNRKPKWK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL 25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	66.0	39	Q9XT68	Q9xt68 canis famil
2	68	66.0	44	Q9GR21	Q9gr21 nematostell
3	68	66.0	60	Q26375	Q26375 styela clav
4	68	66.0	60	Q80WH3	Q80wh3 rattus sp.
5	68	66.0	60	Q80WH2	Q80wh2 rattus sp.
6	68	66.0	60	Q9PS21	Q9ps21 petromyzon
7	68	66.0	60	Q8QGL9	Q8qgl9 petromyzon
8	68	66.0	60	Q91769	Q91769 xenopus lae
9	68	66.0	60	Q9PS22	Q9ps22 petromyzon
10	68	66.0	69	Q9PVR3	Q9pvr3 oryzias lat
11	68	66.0	74	Q57369	Q57369 brachydanio
12	68	66.0	81	Q9PVS5	Q9pvs5 oryzias lat
13	68	66.0	82	Q9PVR4	Q9pvr4 oryzias lat
14	68	66.0	84	Q57363	Q57363 brachydanio
15	68	66.0	85	Q9GP50	Q9gp50 discocella
16	68	66.0	85	Q63255	Q63255 rattus norv

17	68	66.0	85	11	Q61681	Q61681 mus musculu
18	68	66.0	86	13	Q57360	Q57360 brachydanio
19	68	66.0	89	13	Q42371	Q42371 brachydanio
20	68	66.0	93	13	Q801B3	Q801b3 latimeria m
21	68	66.0	103	13	Q93286	Q93286 ambystoma m
22	68	66.0	104	13	Q9PVM1	Q9pvm1 brachydanio
23	68	66.0	106	13	Q801D2	Q801d2 latimeria m
24	68	66.0	125	5	Q96896	Q96896 sacculina c
25	68	66.0	141	11	Q8BQA3	Q8bqa3 mus musculu
26	68	66.0	155	5	Q25577	Q25577 nematostell
27	68	66.0	176	11	Q8BNV9	Q8bnv9 mus musculu
28	68	66.0	181	5	Q26611	Q26611 styela plic
29	68	66.0	182	5	Q9U8Q5	Q9u8q5 dicyema ori
30	68	66.0	185	11	Q8BNX6	Q8bnx6 mus musculu
31	68	66.0	193	5	Q817D0	Q817d0 ciona intes
32	68	66.0	194	5	Q26478	Q26478 styela clav
33	68	66.0	218	11	Q9CUH8	Q9cul8 mus musculu
34	68	66.0	241	13	Q9YH27	Q9yh27 gallus gall
35	68	66.0	242	13	Q9YH13	Q9yh13 gallus gall
36	68	66.0	243	4	Q8N8T3	Q8n8t3 homo sapien
37	68	66.0	244	13	Q9PWL5	Q9pwl5 brachydanio
38	68	66.0	245	13	Q8AWZ0	Q8awz0 brachydanio
39	68	66.0	247	13	Q8JH55	Q8jh55 brachydanio
40	68	66.0	272	13	Q9PSW4	Q9psw4 petromyzon
41	68	66.0	285	11	Q8BP26	Q8bpe6 mus musculu
42	68	66.0	289	4	Q8IXZ1	Q8ixz1 homo sapien
43	68	66.0	363	5	Q9XY00	Q9xy00 dugesia jap
44	68	66.0	392	13	Q42503	Q42503 fugu rubrip
45	68	66.0	435	5	Q817C8	Q817c8 ciona intes
46	67	65.0	33	5	Q86FU0	Q86fu0 drosophila
47	67	65.0	39	13	Q57368	Q57368 brachydanio
48	67	65.0	42	11	Q80WH6	Q80wh6 rattus sp.
49	67	65.0	43	13	Q57359	Q57359 brachydanio
50	67	65.0	46	13	Q9PVR9	Q9pvr9 oryzias lat
51	67	65.0	51	5	Q27413	Q27413 ctenodrilus
52	67	65.0	51	5	Q23743	Q23743 ctenodrilus
53	67	65.0	51	5	Q26407	Q26407 ctenodrilus
54	67	65.0	57	13	Q9PVR8	Q9pvr8 oryzias lat
55	67	65.0	58	5	Q91188	Q9y188 priapulid c
56	67	65.0	58	5	Q25208	Q25208 junonia coe
57	67	65.0	58	13	Q57362	Q57362 brachydanio
58	67	65.0	59	5	Q8WRM9	Q8wr9 lithobius a
59	67	65.0	59	5	Q9NB42	Q9nb42 anopheles g
60	67	65.0	59	13	Q9PVR5	Q9pvr5 oryzias lat
61	67	65.0	60	5	Q77143	Q77143 archegozete
62	67	65.0	60	5	Q77139	Q77139 archegozete
63	67	65.0	60	11	Q80WH7	Q80wh7 rattus sp.
64	67	65.0	60	11	Q80WH4	Q80wh4 rattus sp.
65	67	65.0	60	13	Q8QGL5	Q8qgl5 petromyzon
66	67	65.0	60	13	Q8QGL3	Q8qgl3 petromyzon
67	67	65.0	60	13	Q8QGL6	Q8qgl6 petromyzon
68	67	65.0	60	13	Q8QGL2	Q8qgl2 petromyzon
69	67	65.0	60	13	Q8QGL8	Q8qgl8 petromyzon
70	67	65.0	60	13	Q8QGL7	Q8qgl7 petromyzon
71	67	65.0	61	5	Q27910	Q27910 polyandroca
72	67	65.0	63	5	Q77138	Q77138 archegozete
73	67	65.0	63	5	Q8MX32	Q8mxb2 holopneuste
74	67	65.0	66	13	Q57356	Q57356 brachydanio
75	67	65.0	69	5	Q9U9T4	Q9u9t4 nereis vire
76	67	65.0	69	5	Q9BMF7	Q9bmf7 haliotis as
77	67	65.0	70	5	Q967W5	Q967w5 foliosmia ca
78	67	65.0	70	13	Q801B4	Q801b4 latimeria m
79	67	65.0	71	13	Q9PVS3	Q9pvs3 oryzias lat
80	67	65.0	71	13	Q9PVS1	Q9pvs1 oryzias lat
81	67	65.0	73	5	Q9Y186	Q9y186 priapulid c
82	67	65.0	73	5	Q86D93	Q86d93 spadella ce
83	67	65.0	74	13	Q57367	Q57367 brachydanio
84	67	65.0	75	5	Q25209	Q25209 junonia coe
85	67	65.0	75	13	Q9PVR6	Q9pvr6 oryzias lat
86	67	65.0	76	5	Q44257	Q44257 ethmostigm
87	67	65.0	76	5	Q86NB1	Q86nb1 callanus hel
88	67	65.0	77	5	Q44260	Q44260 ethmostigm
89	67	65.0	77	5	Q9Y187	Q9y187 priapulid c

90 Q9U9Z4 lingula ung
 Q720F3 symagittif
 Q967V2 lithobius f
 Q9U9T9 nereis vire
 Q05008 artemia san
 Q9BN27 porcellio s
 P91769 manduca sex
 Q17142 brachiosteo
 Q9PVR7 oryzias lat
 Q24758 drosophila
 Q9PVS0 oryzias lat

ALIGNMENTS

RESULT 1
 Q9XT68 PRELIMINARY; PRT; 39 AA.
 AC Q9XT68;
 DT 01-NOV-1999 (TREMELrel. 12, Created)
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Hoxa (fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99265967; PubMed=10331940;
 RA Li R., Mignot E., Faraco J., Kadorani H., Cantanese J., Zhao B.,
 RA Lin X., Hinton L., Ostrander E.A., Patterson D.F., de Jong P.J.;
 RT "Construction and characterization of an eightfold redundant dog
 RT genomic bacterial artificial chromosome library."
 RL Genomics 58:9-17(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kodatani H., Mignot E.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF03746; A040572.1; -!
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 39
 SQ SEQUENCE 39 AA; 4908 MW; 1252D75EA245BCAB CRC64;

Query Match 66.0%; Score 68; DB 6; Length 39;
 Best Local Similarity 75.0%; Pred. No. 0.0028;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RPKIWFNNRRKPKWK 17
 Db 13 RQVKIWFNNRRKPKWK 28

RESULT 2
 Q9GR21 PRELIMINARY; PRT; 44 AA.
 AC Q9GR21;
 DT 01-MAR-2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Hox type homeodomain protein (Fragment).
 GN ANTHOXIA.
 OS Nemastostella vectensis.
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Edwardsiidae; Nemastostella.
 OX NCBI_TaxID=45351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97435515; PubMed=9290214;
 RA Finnerty J.R., Martindale M.Q.;
 RT "Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of
 RT Nemastostella vectensis and Metridium senile.";
 RL Biol. Bull. 193:62-76(1997).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF020955; AAG37788.1; -!
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 44
 SQ SEQUENCE 44 AA; 5362 MW; FAF8044070A9CCF3 CRC64;
 Query Match 66.0%; Score 68; DB 5; Length 44;
 Best Local Similarity 75.0%; Pred. No. 0.0031;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 RPKIWFNNRRKPKWK 17
 Db 11 RQVKIWFNNRRKPKWK 26

RESULT 3
 Q26375 PRELIMINARY; PRT; 60 AA.
 AC Q26375;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE DNA binding protein Ahox2 (fragment).
 GN AHOX2.
 OS Styela clava (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Styelidae; Styela.
 OX NCBI_TaxID=7725;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95011617; PubMed=7926803;
 RA Ge T., Lee H., Tomlinson C.R.;
 RT "Identification of an antenapedia-like homeobox gene in the ascidians
 RT Styela clava and S. plicata.";
 RL Gene 147:219-222(1994).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; S73920; AAB33061.2; -!
 DR FIR; PC2399; PC2399.
 DR HSSP; P02833; 9ANT.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.

```

DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7630 MW; F506301E9679BA25 CRC64;

Query Match
Best Local Similarity 66.0%; Score 68; DB 5; Length 60;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKWFPNRRKPKWK 17
Db 43 RQVKWFPNRRMKWK 58

RESULT 4
Q80WH3 PRELIMINARY; PRT; 60 AA.
AC Q80WH3
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hox-B|Hox-2 (Fragment).
GN Hox-B|Hox-2.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10118;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RL Biochem. Genet. 32:351-360(1994).
DR EMBL; S76300; AAP31868.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7671 MW; DD27AFB608CF1F58 CRC64;

Query Match
Best Local Similarity 66.0%; Score 68; DB 11; Length 60;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKWFPNRRKPKWK 17
Db 43 RQVKWFPNRRMKWK 58

RESULT 5
Q80WH2 PRELIMINARY; PRT; 60 AA.
AC Q80WH2
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hox-C|Hox-3 (Fragment).
GN Hox-C|Hox-3.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10118;

```

```

RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=95217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RL Biochem. Genet. 32:351-360(1994).
DR EMBL; S76301; AAP31869.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7602 MW; DD26AEB609CE1F58 CRC64;

Query Match
Best Local Similarity 66.0%; Score 68; DB 11; Length 60;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKWFPNRRKPKWK 17
Db 43 RQVKWFPNRRMKWK 58

RESULT 6
Q9PSZ1 PRELIMINARY; PRT; 60 AA.
AC Q9PSZ1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox protein HoxQ8a (Fragment).
GN HoxQ8a.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=98100508; PubMed=9437854;
RA Carr J.L., Shashikant C.S., Bailey W.J., Ruddle F.H.;
RT "Molecular evolution of Hox gene regulation: cloning and transgenic
analysis of the lamprey HoxQ8 gene.";
RL J. Exp. Zool. 280:73-85(1998).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF035589; AAC04331.1; -
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMO0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
FT NON TER 60
SQ SEQUENCE 60 AA; 7655 MW; DD30A4DBA91F1F58 CRC64;

Query Match
Best Local Similarity 66.0%; Score 68; DB 13; Length 60;

```

```

Best Local Similarity 75.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
DB 43 RQVKIWFQNNRRKWK 58

RESULT 7
Q9QGL9 Q8QGL9 PRELIMINARY; PRT; 60 AA.
AC Q8QGL9;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE HoxF5 homeobox (Fragment).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
RA Ameniya C.T., Ruddle F.H.;
RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
RT marinus."
RL J. Exp. Zool. 0:0-0(2002).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF410910; AAM19468.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7673 MW; 4C2BF6832D53EFA0 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
DB 43 RQVKIWFQNNRRKWK 58

RESULT 8
Q91769 Q81769 PRELIMINARY; PRT; 60 AA.
AC Q91769;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Homeodomain protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Bittner D.A., De Robertis E.M., Cho K.W.;
RX MEDLINE=93326800; PubMed=7916675;

```

```

RT "Characterization of the Xenopus Hox 2.4 gene and identification of
RT control elements in its intron."
RL Dev. Dyn. 196.11-24(1993).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; M83947; AAA49748.1; -.
DR PIR; I51437; I51437.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7671 MW; DD27AFB608CF1F58 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
DB 43 RQVKIWFQNNRRKWK 58

RESULT 9
Q9PSZ2 Q9PSZ2 PRELIMINARY; PRT; 60 AA.
AC Q9PSZ2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Homeobox protein HoxR8 (Fragment).
GN HOXR8.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98100508; PubMed=9437854;
CA Carr J.B., Shashkant C.S., Bailey W.J., Ruddle F.H.;
RT "Molecular evolution of Hox gene regulation: cloning and transgenic
RT analysis of the lamprey Hox8 gene."
RL J. Exp. Zool. 280:73-85(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF035588; AAC04330.1; -.
DR HSSP; P02833; 2HOA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7685 MW; 2C40AFB3089F1F40 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 60;

```



```

Best Local Similarity 75.0%; Pred. No. 0.0042; Mismatches 0; Gaps 0;
Matches 12; Conservative 0; Indels 4;

QY 2 RQPKIWFNNRRKPKWK 17
Db 43 RQVKIWFQNRKWK 58

RESULT 10
Q9PVR3 PRELIMINARY; PRT; 69 AA.
AC Q9PVR3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE HOXB8B (Fragment).
GN HOXB8B.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RT "Hox genes of the medakafish Oryzias latipes."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB026966; BAA86249.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox_1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 69 AA; 8517 MW; C627007BD435E847 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 69;
Best Local Similarity 75.0%; Pred. No. 0.0048;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
Db 23 RQVKIWFQNRKWK 38

RESULT 11
O57369 PRELIMINARY; PRT; 74 AA.
AC O57369;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hoxc8 protein (Fragment).
GN Hoxc8A OR HOXC8.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Prince V.E., Joly L., Ekker M., Ho R.K.;
RT "Zebrafish hox genes: genomic organization and modified colinear

```

```

RT expression patterns in the trunk."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; Y14544; CAA74879.1; -
DR TRANSFAC; T03643; -
DR ZFIN; ZDB-GENE-990415-114; hoxc8a.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 74 AA; 8792 MW; 91D53D3B66C363DD CRC64;

Query Match 66.0%; Score 68; DB 13; Length 74;
Best Local Similarity 75.0%; Pred. No. 0.0052;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17
Db 15 RQVKIWFQNRKWK 30

RESULT 12
Q9PVS5 PRELIMINARY; PRT; 81 AA.
AC Q9PVS5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE HOXA4A (Fragment).
GN HOXA4A.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RT "Hox genes of the medakafish Oryzias latipes."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB026954; BAA86237.1; -
DR HSSP; P02833; 1HOM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 81 AA; 9492 MW; EB28D279F04C686B CRC64;

Query Match 66.0%; Score 68; DB 13; Length 81;
Best Local Similarity 75.0%; Pred. No. 0.0057;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNNRRKPKWK 17

```

Db 23 RQVKIWFQNRMRKWK 38

RESULT 13

Q9PVR4 PRELIMINARY; PRT; 82 AA.

ID Q9PVR4

AC Q9PVR4; 01-MAY-2000 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE HOXC8A (Fragment).

GN HOXC8A.

OS Oryzias latipes (Medaka fish) (Japanese ricefish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

NCBI_TaxID=8090;

[1]

RN SEQUENCE FROM N.A.

RP Kondo S., Naruse K., Shima A.;

RT "Hox genes of the medakafish Oryzias latipes.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AB026965; BAA86248.1; -;

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006335; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR000047; HTH_lambrepreser.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR PRINTS; PR00031; HTHREPRESS.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SMO0389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX_1; 1.

DR PROSITE; PS00711; HOMEBOX_2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON-TER

SQ SEQUENCE 82 AA; 9782 MW; 5BF76FE195796069 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 82;

Best Local Similarity 75.0%; Pred. No. 0.0057;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQVKIWFQNRMRKWK 17

DB 23 RQVKIWFQNRMRKWK 38

RESULT 14

Q57363 PRELIMINARY; PRT; 84 AA.

ID Q57363

AC Q57363; 01-JUN-1998 (TREMELrel. 06, Created)

DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Hoxb8 protein (Fragment).

GN HOXB8A OR HOXB8.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

NCBI_TaxID=7955;

[1]

RN SEQUENCE FROM N.A.

RP Prince V.E., Joly L., Ekker M., Ho R.K.;

RT "zebrafish hox genes: genomic organization and modified colinear

RT expression patterns in the trunk.";

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; Y14530; CAA74865.1; -;

A;Accession: PC2400
A;Molecule type: DNA
A;Residues: 1-60 <GEN>
A;Note: The authors translated the codon ATA for residue 47 as Glu
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0082;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 43 RQVKIWFQNRKMKWK 58

RESULT 10
I51437
homeotic protein Hox 2.4 - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51437
R;Bittner, D.
Dev. Dyn. 196, 11-24, 1993
A;Title: Characterization of the Xenopus Hox 2.4 gene and identification of control elements
A;Reference number: I51437; MUID:93326800; PMID:7916675
A;Accession: I51437
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-60 <BIT>
A;Cross-references: GB:M83947; NID:G214252; PIDN:AAA49748.1; PID:G214253
C;Genetics:
A;Gene: Hox 2.4
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0082;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 43 RQVKIWFQNRKMKWK 58

RESULT 11
B32391
homeotic protein Hox 2.4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Dec-1989 #sequence_revision 08-Dec-1989 #text_change 17-Oct-1997
C;Accession: B32391
R;Graham, A.; Papalopulu, N.; Krumlauf, R.
Cell 57, 367-378, 1989
A;Title: The murine and Drosophila homeobox gene complexes have common features of organogenesis
A;Reference number: A32391; MUID:89249299; PMID:2566383
A;Accession: B32391
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-62 <GRA>
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 62;
Best Local Similarity 75.0%; Pred. No. 0.0084;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 43 RQVKIWFQNRKMKWK 58

RESULT 12
S15534
homeotic protein Hox C8 - human (fragment)
N;Alternate names: homeotic protein Hox 3A
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
C;Accession: S15534
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Stoc
Genome 31, 745-756, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; MUID:90215256; PMID:2576652
A;Accession: S15534
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-66 <BON>
C;Genetics:
A;Gene: GDB:H0XC8
A;Cross-references: GDB:I20668; OMIM:142970
A;Map position: 12q13.3-12q13.3
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 66;
Best Local Similarity 75.0%; Pred. No. 0.0088;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 43 RQVKIWFQNRKMKWK 58

RESULT 13
B27471
homeotic protein R2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 17-Nov-2000
C;Accession: B27471; I65243
R;Falzon, M.; Sanderson, N.; Chung, S.Y.
Gene 54, 23-32, 1987
A;Title: Cloning and expression of rat homeo-box-containing sequences.
A;Reference number: A91576; MUID:87277429; PMID:2886401
A;Accession: B27471
A;Molecule type: DNA
A;Residues: 1-67 <FAL>
R;Sakoyama, Y.; Mizuta, I.; Ogasawara, N.; Yoshikawa, H.
Biochem. Genet. 32, 351-360, 1994
A;Title: Cloning of rat homeobox genes.
A;Reference number: I52340; MUID:95217128; PMID:7702549
A;Accession: I65243
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-42 <RES>
A;Cross-references: GB:S76296; NID:G913079
C;Genetics:
A;Gene: Hox-A; Hox-1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 52.7%; Score 68; DB 2; Length 67;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 25 RQVKIWFQNRKMKWK 40

RESULT 14
S13785
homeotic protein m31 - mouse (fragment)

Query Match 52.7%; Score 68; DB 2; Length 85;
 Best Local Similarity 75.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
 |||||
 Db 52 RQVKIWFQNRMRKWK 67

RESULT 19

homeotic protein Hox D8 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 29-Jan-1990 #sequence_revision 09-Mar-1996 #text_change 24-Sep-1999
 C:Accession: B32830; S14937; S05957
 R:Oliver, G.; Sidell, N.; Fiske, W.; Heinzmann, C.; Mohandas, T.; Sparkes, R.S.; De Robertis, G.
 A:Title: Complementary homeo protein gradients in developing limb buds.
 A:Reference number: A32830; MUID:89306602; PMID:2568311
 A:Accession: B32830
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-95
 A:Cross-references: EMBL:X15507; NID:g32399; PIDN:CAA33529.1; PID:g32400
 R:Acampora, D.; d'Esposito, M.; Faella, A.; Pannese, M.; Migliaccio, E.; Morelli, F.; S. Nucleic Acids Res. 17, 10385-10402, 1989
 A:Title: The human HOX gene family.
 A:Reference number: S07541; MUID:90098876; PMID:2574852
 A:Accession: S14937
 A:Molecule type: DNA
 A:Residues: 5-70 <ACA>
 C:Genetics:
 A:Gene: GDB:HOXD8
 A:Cross-references: GDB:120680; OMIM:142985
 A:Map position: 2q31-2q31
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:6-62/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 95;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
 |||||
 Db 47 RQVKIWFQNRMRKWK 62

RESULT 20

homeotic protein HOX3A - cat (fragment)
 C:Species: Felis silvestris catus (domestic cat)
 C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Oct-1997
 C:Accession: A55278
 R:Masuda, R.; Yuhki, N.; O'Brien, S.J.
 Genomics 11, 1007-1013, 1991
 A:Title: Molecular cloning, chromosomal assignment, and nucleotide sequence of the feline homeobox gene HOX3A.
 A:Reference number: A55278; MUID:92147095; PMID:1686012
 A:Accession: A55278
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-96 <NAS>
 A:Note: sequence extracted from NCBI backbone (NCBIN:80214, NCBI:80215)
 C:Genetics:
 A:Gene: HOX3A
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:4-60/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 96;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
 |||||
 Db 45 RQVKIWFQNRMRKWK 60

RESULT 21

homeotic protein Hox 4.3 - mouse (fragment)
 A:1605
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1992 #sequence_revision 01-Dec-1995 #text_change 24-Sep-1999
 C:Accession: S16177; C42694; A41605; A38810; S15521
 R:Sadoul, R.; Featherstone, M.S.
 Biochim. Biophys. Acta 1089, 259-261, 1991
 A:Title: Sequence analysis of the homeobox-containing exon of the murine Hox-4.3 homeo gene.
 A:Reference number: S16177; MUID:91274361; PMID:1675873
 A:Accession: S16177
 A:Molecule type: DNA
 A:Residues: 1-99 <BIO>
 A:Cross-references: EMBL:X56561; NID:G51418; PIDN:CAA39911.1; PID:e30697; PID:g1333935
 R:Nazarali, A.; Kim, Y.; Nirenberg, M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2883-2887, 1992
 A:Title: Hox-1.11 and Hox-4.9 homeobox genes.
 A:Reference number: A42694; MUID:92212934; PMID:1348361
 A:Accession: C42694
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 2-70 <NAZ>
 A:Note: sequence extracted from NCBI backbone (NCBIN:92289, NCBI:92290)
 R:Singh, G.; Kaur, S.; Stock, J.L.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Potter
 Proc. Natl. Acad. Sci. U.S.A. 88, 10706-10710, 1991
 A:Title: Identification of 10 murine homeobox genes.
 A:Reference number: A37290; MUID:92073356; PMID:1683707
 A:Accession: A41605
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 5-64 <SI2>
 C:Genetics:
 A:Gene: Hox 4.3
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:6-62/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 99;
 Best Local Similarity 75.0%; Pred. No. 0.012;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
 |||||
 Db 47 RQVKIWFQNRMRKWK 62

RESULT 22

homeotic protein R4 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 24-Sep-1999
 C:Accession: D43559
 R:Falzon, M.; Chung, S.Y.
 Development 103, 601-610, 1988
 A:Title: The expression of rat homeobox-containing genes is developmentally regulated at the embryonic level.
 A:Reference number: A43559; MUID:89231502; PMID:2907739
 A:Accession: D43559
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-104 <FAL>
 A:Cross-references: GB:M37568; NID:g204636; PIDN:AAA41344.1; PID:g204637
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:12-68/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 104;

```
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRRKPWK 21
Db 53 RQVKIWFQNRMRMKWK 68

RESULT 23
A43559
homeotic protein Rla - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1992 #sequence_revision 30-Jan-1993 #text_change 24-Sep-1999
C;Accession: A43559
R;Falzon, M.; Chung, S.Y.
Development 103, 601-610, 1988
A;Title: The expression of rat homeobox-containing genes is developmentally regulated and
A;Reference number: A43559; MUID:89231502; PMID:2907739
A;Accession: A43559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <FAL>
A;Cross-references: GB:M37565; NID:G204628; PIDN:AAA41341.1; PID:G204629
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;12-68/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 114;
Best Local Similarity 75.0%; Pred. No. 0.014;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRRKPWK 21
Db 53 RQVKIWFQNRMRMKWK 68

RESULT 24
S25846
homeotic protein Hox A4, testicular - mouse (fragment)
N;Alternate names: homeotic protein Hox 1.4
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 06-Feb-1998 #text_change 24-Sep-1999
C;Accession: S25847; S25846
R;Rubin, M.R.; Nguyen-Huu, M.C.
DNA Seq. 1, 329-334, 1991
A;Title: Murine embryonic spinal cord and adult testis Hox-1.4 cDNAs are identical 3' to
A;Reference number: S25846; MUID:92190549; PMID:1686835
A;Accession: S25847
A;Molecule type: DNA
A;Residues: 1-28 <RUB1>
A;Cross-references: EMBL:X17346; NID:G51375; PID:G51376
A;Accession: S25846
A;Molecule type: mRNA
A;Residues: 27-155 <RUB2>
A;Cross-references: EMBL:X13538; NID:G51373; PIDN:CAA31889.1; PID:G51374
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;10-65/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 155;
Best Local Similarity 75.0%; Pred. No. 0.018;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRRKPWK 21
Db 51 RQVKIWFQNRMRMKWK 66

RESULT 25
I50145
homeotic protein Hox M - chicken
N;Alternate names: CHOX M
C;Species: Gallus gallus (chicken)
```

```
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I50145; S14512
R;Crompton, M.R.; MacGregor, A.D.; Goodwin, G.H.
Leukemia 5, 357-360, 1991
A;Title: cDNA cloning of a homeobox-containing gene expressed in avian myeloblastic virus
A;Reference number: I50145; MUID:91238215; PMID:1674560
A;Accession: I50145
A;Status: preliminary; translated from GB/EMBL/DDEJ
A;Molecule type: mRNA
A;Residues: 1-188 <CRO>
A;Cross-references: EMBL:X57158; NID:G62700; PIDN:CAA40445.1; PID:G62701
C;Genetics:
A;Gene: CHOX M
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;96-152/Domain: homeobox homology <HOX>

Query Match 52.7%; Score 68; DB 2; Length 188;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRRKPWK 21
Db 137 RQVKIWFQNRMRMKWK 152

Search completed: May 24, 2004, 17:26:11
Job time : 12.2027 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:17:30 ; Search time 7.66216 Seconds
(without alignments)
142.711 Million cell updates/sec

Title: US-09-977-349-4

Perfect score: 129

Sequence: 1 CSSCIRQPKIWFNRRKPKK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	55.8	234	1 HXC6 NOTVI	P14858 notophthalm
2	69	53.5	48	1 HXC6 XENLA	P11256 xenopus lae
3	69	53.5	76	1 HXB6 BRARE	P18865 rattus norv
4	69	53.5	84	1 HXB6 CHICK	P14839 gallus gall
5	69	53.5	245	1 HXB4 CHICK	P14840 gallus gall
6	69	53.5	264	1 HXC4 HUMAN	P09017 homo sapien
7	69	53.5	264	1 HXC4 MOUSE	Q08624 mus musculu
8	68	52.7	67	1 HXA4 RAT	P09635 rattus norv
9	68	52.7	92	1 HXB8 RAT	P18863 rattus norv
10	68	52.7	108	1 HXC8 RAT	P18866 rattus norv
11	68	52.7	188	1 HXD8 CHICK	P23459 gallus gall
12	68	52.7	240	1 HXD8 HETFR	Q91a12 heterodontu
13	68	52.7	242	1 HXC8 HUMAN	P09025 mus musculu
14	68	52.7	243	1 HXC8 MOUSE	P09632 mus sapien
15	68	52.7	243	1 HXB8 HUMAN	P17481 homo sapien
16	68	52.7	243	1 HXB8 MOUSE	Q91a22 heterodontu
17	68	52.7	243	1 HXA4 HETFR	Q91a22 heterodontu
18	68	52.7	248	1 HXA4 MORSA	Q91a22 heterodontu
19	68	52.7	285	1 HXA4 MOUSE	Q91a22 heterodontu
20	68	52.7	289	1 HXD8 MOUSE	P06798 mus musculu
21	68	52.7	290	1 HXD8 HUMAN	P23463 mus musculu
22	68	52.7	309	1 HXA4 CHICK	P13378 homo sapien
23	68	52.7	320	1 HXA4 HUMAN	P17277 gallus gall
24	67	51.9	49	1 HXA4 HUMAN	Q00056 homo sapien
25	67	51.9	71	1 HXA7 SHEEP	Q28599 ovis aries
26	67	51.9	71	1 HXC5 NOTVI	Q28600 ovis aries
27	67	51.9	74	1 HXC5 APIME	P15860 apis mellif
28	67	51.9	75	1 HXA5 SALSA	P15860 apis mellif
29	67	51.9	78	1 HXA5 SALSA	P09636 salmo salar
30	67	51.9	80	1 HXA4 LINS	P09637 salmo salar
31	67	51.9	81	1 HXA4 LINS	P81192 lineus sang
32	67	51.9	82	1 HXB5 BRARE	P09013 brachydario
33	67	51.9	86	1 SCR APIME	P15859 apis mellif

34	67	51.9	87	1 HXC5 XENLA	P09020 xenopus lae
35	67	51.9	93	1 HXB8 PIG	P09078 sus scrofa
36	67	51.9	96	1 HXC6 BRARE	P15862 brachydario
37	67	51.9	105	1 HXA7 RAT	P09634 rattus norv
38	67	51.9	105	1 HXB4 BRARE	P22574 brachydario
39	67	51.9	112	1 HXB7 RAT	P18864 rattus norv
40	67	51.9	148	1 HXA5 AMEME	P50208 ambystoma m
41	67	51.9	153	1 HXC6 SHEEP	P49925 ovis aries
42	67	51.9	208	1 HXA7 HETFR	Q91a25 heterodontu
43	67	51.9	209	1 HXA7 XENLA	P09071 xenopus lae
44	67	51.9	217	1 HXB7 BOVIN	Q91a25 heterodontu
45	67	51.9	217	1 HXB7 HUMAN	Q91a25 heterodontu
46	67	51.9	217	1 HXB7 MOUSE	Q91a25 heterodontu
47	67	51.9	220	1 HXB7A XENLA	P09024 mus musculu
48	67	51.9	220	1 HXB7B XENLA	Q91a25 heterodontu
49	67	51.9	222	1 HXC5 HUMAN	Q00444 homo sapien
50	67	51.9	222	1 HXC5 MOUSE	P12043 mus musculu
51	67	51.9	224	1 HXB6 HUMAN	P17509 homo sapien
52	67	51.9	224	1 HXB6 MOUSE	P09023 mus musculu
53	67	51.9	225	1 HXA7 MORSA	Q91a25 heterodontu
54	67	51.9	228	1 HXB6 BRARE	P15861 brachydario
55	67	51.9	229	1 HXA6 HETFR	Q91a24 heterodontu
56	67	51.9	229	1 HXA7 MOUSE	P02830 mus musculu
57	67	51.9	230	1 HXA7 HUMAN	P31268 homo sapien
58	67	51.9	230	1 HXB5 XENLA	P09019 xenopus lae
59	67	51.9	232	1 HXA6 MOUSE	P09092 mus musculu
60	67	51.9	232	1 HXB4 XENLA	P09070 xenopus lae
61	67	51.9	232	1 HXC5 BRARE	P09074 brachydario
62	67	51.9	233	1 HXC5 RAT	P52949 rattus norv
63	67	51.9	233	1 HXA6 HUMAN	P31267 homo sapien
64	67	51.9	234	1 HXC6 XENLA	P02832 xenopus lae
65	67	51.9	235	1 HXC6 HUMAN	P09630 homo sapien
66	67	51.9	235	1 HXC6 MOUSE	P10629 mus musculu
67	67	51.9	235	1 HXD4 CHICK	P17278 gallus gall
68	67	51.9	236	1 HXD4 BRARE	O57374 brachydario
69	67	51.9	242	1 HXA7 COTUA	P24061 coturnix co
70	67	51.9	250	1 HXB4 MOUSE	P10284 mus musculu
71	67	51.9	250	1 HXD4 MOUSE	P10628 mus musculu
72	67	51.9	251	1 HXB4 FUGRU	O13074 fugu rubrip
73	67	51.9	251	1 HXB4 HUMAN	P17483 homo sapien
74	67	51.9	252	1 HXD5 HETFR	Q91a11 heterodontu
75	67	51.9	255	1 HXD4 HUMAN	P09016 homo sapien
76	67	51.9	261	1 HXC4 ORYLA	Q91a11 heterodontu
77	67	51.9	269	1 HXB5 HUMAN	Q91a11 heterodontu
78	67	51.9	269	1 HXB5 MOUSE	Q91a11 heterodontu
79	67	51.9	270	1 HXB5 HUMAN	P09067 homo sapien
80	67	51.9	270	1 HXA5 HUMAN	P09079 mus musculu
81	67	51.9	270	1 HXA5 MOUSE	P20719 homo sapien
82	67	51.9	275	1 HXA5 HETFR	P09021 mus musculu
83	67	51.9	275	1 HXB5 BRARE	Q91a23 heterodontu
84	67	51.9	281	1 HXA5 MORSA	Q91a23 heterodontu
85	67	51.9	378	1 HMAN DROME	Q91a23 heterodontu
86	67	51.9	394	1 HMAN DROSU	Q91a23 heterodontu
87	67	51.9	415	1 SCR DROME	Q91a23 heterodontu
88	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
89	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
90	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
91	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
92	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
93	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
94	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
95	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
96	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
97	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
98	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
99	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu
100	67	51.9	590	1 HMD F DROME	Q91a23 heterodontu

ALIGNMENTS

RESULT 1

RESULT 2

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X17612; CAA35614.1; --
PIR; S10092; S10092.
HSSP; P02833; 9ANT.
TRANSFAC; T01726; --
InterPro; IPR001827; Antennapedia.
InterPro; IPR001356; Homeobox.
Pfam; PF00046; homeobox; 1.
PRINTS; PR00025; ANTENNAPEDIA.
PRINTS; PR00024; HOMEBOX.
ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEBOX 1; 1.
PROSITE; PS00071; HOMEBOX 2; 1.
PROSITE; PS00032; ANTENNAPEDIA; 1.
Homeobox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation.
SITE 129 134 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 150 209 HOMEBOX.
SQ SEQUENCE 245 AA; 27782 MW; 4949B200FEC44E91 CRC64;

Query Match 53.5%; Score 69; DB 1; Length 245;
Best Local Similarity 63.6%; Pred. No. 0.0033;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 2 SSC1--RQPKIWFNRRKPKWK 21
| : | | | | | | | | | |
Db 186 SLCLSERQIKIWFQNRMRKWK 207

RESULT 6

HXC4 HUMAN
ID HXC4 HUMAN STANDARD; PRT; 264 AA.
AC P09017;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-C4 (Hox-3E) (CP19).
GN HXC4 OR HOXC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88262550; PubMed=2898768;
RA Simeone A., Pannese M., Acampora D., D'Esposito M., Boncinelli E.;
RT "At least three human homeoboxes on chromosome 12 belong to the same
RT transcription unit.";
RL Nucleic Acids Res. 16:5379-5390(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.;
RT "A complete mutation analysis panel of human HOX genes.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 156-221 FROM N.A.
RX MEDLINE=90215256; PubMed=2576652;
RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RL Genome 31:745-756(1989).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X07495; CAA30376.1; --
EMBL; AY014298; AAG42145.1; --
EMBL; AY014297; AAG42145.1; JOINED.
PIR; S01030; WJHU3E.
HSSP; P02833; 9ANT.
TRANSFAC; T03325; --
Genew; HGNC:5126; HOXC4.
MIM; 142974; --
GO; GO:0005634; C:nucleus; NAS.
GO; GO:0003700; P:transcription factor activity; NAS.
GO; GO:0007275; P:development; NAS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
InterPro; IPR001827; Antennapedia.
InterPro; IPR001356; Homeobox.
Pfam; PF00046; homeobox; 1.
PRINTS; PR00025; ANTENNAPEDIA.
ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00032; ANTENNAPEDIA; 1.
PROSITE; PS00027; HOMEBOX 1; 1.
PROSITE; PS00071; HOMEBOX 2; 1.
Homeobox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation.
DOMAIN 54 60 POLY-PRO.
FT SITE 135 140 ANTP-TYPE HEXAPEPTIDE.
FT DNA BIND 156 215 HOMEBOX.
FT DOMAIN 183 186 POLY-ARG.
SQ SEQUENCE 264 AA; 23781 MW; BDCD139955653373 CRC64;

Query Match 53.5%; Score 69; DB 1; Length 264;
Best Local Similarity 63.6%; Pred. No. 0.0035;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 2 SSC1--RQPKIWFNRRKPKWK 21
| : | | | | | | | | | |
Db 192 SLCLSERQIKIWFQNRMRKWK 213

RESULT 7

HXC4 MOUSE
ID HXC4 MOUSE STANDARD; PRT; 264 AA.
AC Q08624;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-C4 (Hox-3.5).
GN HOXC4 OR HOXC-4 OR HOX-3.5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RX MEDLINE=93288004; PubMed=8099712;
RA Goto J., Miyabayashi T., Wakamatsu Y., Takahashi N., Muramatsu M.-A.;
RT "Organization and expression of mouse Hox3 cluster genes.";
RL Mol. Gen. Genet. 239:41-48(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=93161956; PubMed=1363091;
RA Geadia A.M.C., Gaunt S.J., Azzawi M., Shimeld S.M., Pearce J.,

```

RA  Sharpe P.T.;
RL  "Sequence and embryonic expression of the murine Hox-3.5 gene.";
RN  Development 116:497-506(1992).
RC  (3)
RX  SEQUENCE OF 177-201 FROM N.A.
RY  STRAIN=CS7BL/6; TISSUE=Spleen;
RZ  MEDLINE=92073357; PubMed=1720547;
RA  Murtha M.T., Leckman J.F., Ruddle F.H.;
RT  "Detection of homeobox genes in development and evolution.";
RL  Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).
CC  -1- FUNCTION: Sequence-specific transcription factor which is part of
CC  a developmental regulatory system that provides cells with
CC  specific positional identities on the anterior-posterior axis.
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC  subfamily.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D11328; BAA01947.1; -
DR  EMBL; S62287; AAB27153.1; -
DR  EMBL; X69019; CAA48784.1; -
DR  EMBL; M81660; AAA63313.1; -
DR  PIR; S35219; S35219.
DR  HSSP; P02833; 9ANT.
DR  TRANSFAC; T03340; -.
DR  MGD; MGI:96195; Hoxc4.
DR  InterPro; IPR001827; Antennapedia.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00025; ANTENNAPEDIA.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00032; ANTENNAPEDIA; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Transcription regulation.
FT  DOMAIN 54 60 POLY-PRO.
FT  SITE 135 140 ANTP-TYPE HEXAPEPTIDE (BY SIMILARITY).
FT  DNA_BIND 156 215 HOMEBOX (BY SIMILARITY).
FT  DOMAIN 183 186 POLY-ARG.
FT  CONFLICT 80 80 A -> G (IN REF. 2).
FT  CONFLICT 96 96 P -> S (IN REF. 2).
SQ  SEQUENCE 264 AA; 29865 MW; 35B0FA34B45BF30C CRC64;

Query Match 53.5%; Score 69; DB 1; Length 264;
Best Local Similarity 63.6%; Pred. No. 0.0035;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRKPWK 21
DB 192 SLCLSERQIKIWFQNRKWK 213

RESULT 8
HXA4 RAT
ID HXA4 RAT STANDARD; PRT; 67 AA.
AC P09635;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (R2) (Fragment).
GN HXA4 OR HOX-A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC (1)
CC SEQUENCE FROM N.A.
CC STRAIN=Sprague-Dawley;
CC MEDLINE=87277429; PubMed=2886401;
CC Falzon M., Sanderson N., Chung S.Y.;
CC "Cloning and expression of rat homeo-box-containing sequences.";
CC Gene 54:23-32(1987).
CC (2)
CC SEQUENCE OF 1-42 FROM N.A.
CC MEDLINE=95217128; PubMed=7702549;
CC Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
CC "Cloning of rat homeobox genes.";
CC Biochem. Genet. 32:351-360(1994).
CC -1- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; M16808; -; NOT ANNOTATED_CDS.
DR  EMBL; S76296; -; NOT ANNOTATED_CDS.
DR  PIR; B27471; B27471.
DR  HSSP; P02833; 9ANT.
DR  InterPro; IPR001827; Antennapedia.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00071; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Transcription regulation.
FT  NON_TER 1 1
FT  DNA_BIND <1 42 HOMEBOX.
FT  NON_TER 67 67
SQ  SEQUENCE 67 AA; 8346 MW; 5BDFEAB93193780 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 67;
Best Local Similarity 75.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNNRKPWK 21
DB 25 RQVKIWFQNRKWK 40

RESULT 9
HXB8 RAT
ID HXB8 RAT STANDARD; PRT; 92 AA.
AC P18863;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B8 (R1A) (Fragment).
GN HXB8 OR HOXB-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC (1)

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=89231502; PubMed=2907739;
RA Falzon M., Chung S.Y.;
RT "The expression of rat homeobox-containing genes is developmentally
RL regulated and tissue specific.";
RL Development 103:601-610(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M37565; AAA41341.1; ALT_SEQ.
DR PIR; A43559; A43559.
DR HSSP; P02833; 9ANT.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HTH_lambrepresr.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON_TER 1 1
FT DNA_BIND 11 70 HOMEBOX.
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 11358 MW; 3D8E3EA42622B4CA CRC64;

Query Match 52.7%; Score 68; DB 1; Length 92;
Best Local Similarity 75.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 53 RQVKIWFQNRMRKWK 68
||| ||||| ||| |||

RESULT 10
HXCB RAT STANDARD; PRT; 108 AA.
AC P18856;
DT 01-NOV-1990 (Rel. 16, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C8 (R4) (Fragment).
GN HOC8 OR HOCX-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=55217128; PubMed=7702549;
RA Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT "Cloning of rat homeobox genes.";
RL Biochem. Genet. 32:351-360(1994).
RN [2]

```

```

RP SEQUENCE OF 5-108 FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=89231502; PubMed=2907739;
RA Falzon M., Chung S.Y.;
RT "The expression of rat homeobox-containing genes is developmentally
RL regulated and tissue specific.";
RL Development 103:601-610(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; S76301; -; NOT ANNOTATED_CDS.
DR PIR; M37568; AAA41344.1; -.
DR HSSP; P02833; 9ANT.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HTH_lambrepresr.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON_TER 1 1
FT DNA_BIND 15 74 HOMEBOX.
FT NON_TER 108 AA; 13115 MW; DBFC0AD6A0B08E1 CRC64;
SQ SEQUENCE 108 AA; 13115 MW; DBFC0AD6A0B08E1 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 108;
Best Local Similarity 75.0%; Pred. No. 0.0021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQPKIWFPPNRKPKWK 21
Db 57 RQVKIWFQNRMRKWK 72
||| ||||| ||| |||

RESULT 11
HXCB CHICK STANDARD; PRT; 188 AA.
AC P23459;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D8 (Chox-M).
GN HOD8 OR CHOX-M.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91238215; PubMed=1674560;
RX Crompton M.R., McGregor A.D., Goodwin G.H.;
RT "cDNA cloning of a homeobox-containing gene expressed in avian
RL myeloblastic virus-transformed chicken monoblastic leukaemia cells.";
RL Leukemia 5:357-360(1991).

```

```

CC CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC CC a developmental regulatory system that provides cells with
CC CC specific positional identities on the anterior-posterior axis.
CC CC -!- SUBCELLULAR LOCATION: Nuclear.
CC CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; X57158; CAA40445.1; -.
CC CC PIR; I50145; I50145.
CC CC HSSP; P02833; 9ANT.
CC CC TRANSPAC; T01754; -.
CC CC InterPro; IPR001827; Antennapedia.
CC CC InterPro; IPR001356; Homeobox.
CC CC InterPro; IPR000047; HTH_lambdarepressor.
CC CC Pfam; PF00046; homeobox; 1.
CC CC PRINTS; PR00025; ANTENNAPEDIA.
CC CC PRINTS; PR00031; HTHREPRESSR.
CC CC ProDom; PD000010; Homeobox; 1.
CC CC SMART; SM00389; HOX; 1.
CC CC PROSITE; PS00027; HOMEBOX 1; 1.
CC CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC CC PROSITE; PS00071; HOMEBOX 2; 1.
CC CC KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC CC Transcription regulation.
CC CC FT DOMAIN 15 25 POLY-ALA.
CC CC FT SITE 83 88 ANTP-TYPE HEXAPEPTIDE.
CC CC FT DNA BIND 95 154 HOMEBOX.
CC CC SQ SEQUENCE 188 AA; 21729 MW; D4560E8807FE29FE CRC64;

Query Match 52.7%; Score 68; DB 1; Length 188;
Best Local Similarity 75.0%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNNRRKPKWK 21
Db 137 RQVKIWFQNRKMKWK 152

RESULT 12
HXD8_HETFR
ID HXDB_HETFR STANDARD; PRT; 240 AA.
AC Q9IA12;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D8.
GN HOXD8.
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OX NCBI_TaxID=7792;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20144096; PubMed=10677514;
RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,
RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;
RA "Hox cluster genomics in the horn shark, Heterodontus francisci.";
RA Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).
CC CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC CC a developmental regulatory system that provides cells with
CC CC specific positional identities on the anterior-posterior axis (By
CC CC similarity).
CC CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC CC -!- SIMILARITY: Belongs to the Antp homeobox family.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; AF224263; AAF44632.1; -.
CC CC HSSP; P02833; 9ANT.
CC CC TRANSPAC; T04482; -.
CC CC InterPro; IPR001827; Antennapedia.
CC CC InterPro; IPR001356; Homeobox.
CC CC InterPro; IPR000047; HTH_lambdarepressor.
CC CC Pfam; PF00046; homeobox; 1.
CC CC PRINTS; PR00025; ANTENNAPEDIA.
CC CC PRINTS; PR00031; HTHREPRESSR.
CC CC ProDom; PD000010; Homeobox; 1.
CC CC SMART; SM00389; HOX; 1.
CC CC PROSITE; PS00027; HOMEBOX 1; 1.
CC CC PROSITE; PS00071; HOMEBOX 2; 1.
CC CC KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC CC Transcription regulation.
CC CC FT SITE 135 140 ANTP-TYPE HEXAPEPTIDE.
CC CC FT DNA BIND 147 206 HOMEBOX.
CC CC SQ SEQUENCE 240 AA; 28127 MW; 4B9CC002CE58105 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 240;
Best Local Similarity 75.0%; Pred. No. 0.0045;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNNRRKPKWK 21
Db 189 RQVKIWFQNRKMKWK 204

RESULT 13
HXCB_HUMAN
ID HXC8_HUMAN STANDARD; PRT; 242 AA.
AC P31273; O15221; O15362;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein Hox-C8 (Hox-3A).
GN HXC8 OR HOX3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.;
RA "A complete mutation analysis panel of human HOX genes.";
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Roha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[3]
RP SEQUENCE OF 64-114 AND 147-195 FROM N.A.
RX MEDLINE=98019011; PubMed=9357979;
RA Flagiello D., Gibaud A., Dutrillaux B., Poupon M.F., Malfroy B.;
RT "Distinct patterns of all-trans retinoic acid dependent expression of
RT HOXB and HOXC homeoboxes in human embryonal and small-cell lung
RT carcinoma cell lines.";
RL FEBS Lett. 415:263-267 (1997).
[4]
RP SEQUENCE OF 149-214 FROM N.A.
RX MEDLINE=90215256; PubMed=2576652;
RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RL Genome 31:745-756 (1989).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBUNIT: INTERACTS WITH SMAD1.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AY014300; AAG42146.1; -
DR EMBL; AY014299; AAG42146.1; JOINED.
DR EMBL; BC053898; AAH53898.1; -
DR EMBL; X99680; CAA67996.1; -
DR EMBL; X99681; CAA67997.1; -
DR PIR; S15534; S15534.
DR HSPF; P02833; 9ANT.
DR TRANSFAC; T03326; -
DR MIM; 142970; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0007275; P:development; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepres.
DR Pfam; PF000045; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESS.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00027; HOMEBOX; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 138 143 ANT-P-TYPE HEXAPEPTIDE.
FT DNA BIND 149 208 HOMEBOX.
FT CONFLICT 165 165 E -> Q (IN REF. 4).
SQ SEQUENCE 242 AA; 27754 MW; 30C1FD06228833FC CRC64;
Query Match 52.7%; Score 68; DB 1; Length 242;
Best Local Similarity 75.0%; Pred. No. 0.0046;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 6 ROPKIWPNRKPKWK 21
DB 191 RQVKIWPQNRKPKWK 206

RESULT 14
HX88 MOUSE
ID HX88 MOUSE STANDARD; PRT; 242 AA.
AC P09025;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-C8 (Hox-3.1) (M31).
GN HOXC8 OR HOXC-8 OR HOX-3.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88185816; PubMed=2895723;
RA le Mouellie H., Condamine H., Brulet P.;
RT "Pattern of transcription of the homeo gene Hox-3.1 in the mouse
RT embryo.";
RL Genes Dev. 2:125-135 (1988).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88312579; PubMed=2900757;
RA Breier G., Dressler G.R., Gruss P.;
RT "Primary structure and developmental expression pattern of Hox 3.1, a
RT member of the murine Hox 3 homeobox gene cluster.";
RL EMBO J. 7:1329-1336 (1988).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Connective tissue;
RX MEDLINE=90349629; PubMed=1696731;
RA Angulewitsch A., Bieberich C., Bogarad L., Shashikant C., Ruddle F.H.;
RT "Structural analysis of the Hox-3.1 transcription unit and the
RT Hox-3.2-Hox-3.1 intergenic region.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6428-6432 (1990).
[4]
RP SEQUENCE OF 147-216 FROM N.A.
RX MEDLINE=86175026; PubMed=3007994;
RA Angulewitsch A., Utset M.F., Hart C.P., McGinnis W., Ruddle F.H.;
RT "Spatial restriction in expression of a mouse homeo box locus within
RT the central nervous system.";
RL Nature 320:328-335 (1986).
[5]
RP SEQUENCE OF 149-215 FROM N.A.
RX MEDLINE=87053860; PubMed=2877873;
RA Breier G., Bucan M., Francke U., Colberg-Poley A.M., Gruss P.;
RT "Sequential expression of murine homeo box genes during F9 EC cell
RT differentiation.";
RL EMBO J. 5:2209-2215 (1986).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: Initially found in all tissues of the
CC posterior region in 8.5 and 9.5 dpc. Embryos, it eventually become
CC specifically located in neural tissue.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

T	"Organization of human class I homeobox genes."	RC
L	Genome 31:745-756(1989).	RX
		RA

RT "Expression of Hox-2.4 homeobox gene directed by proviral insertion
 in a myeloid leukemia.";
 RL Nucleic Acids Res. 17:1881-1892(1989).
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;
 RX MEDLINE=89210815; PubMed=2907477;
 RA Blatt C., Aberdam D., Schwartz R., Sachs L.;
 RT "DNA rearrangement of a homeobox gene in myeloid leukaemic cells.";
 RL EMBO J. 7:4283-4290(1988).
 RN [3]

RP REVISIONS:

RA Blatt C.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=91220677; PubMed=1673811;
 RA Ben-David L., Aberdam D., Sachs L., Blatt C.;
 RT "A deletion and a rearrangement distinguish between the
 intracisternal A- particle of Hox-2.4 and that of interleukin-3 in
 the same leukemic cells.";
 RL Virology 182:382-387(1991).
 RN [5]

RP SEQUENCE OF 144-243 FROM N.A.

RX MEDLINE=88085193; PubMed=2891508;
 RA Hart C.P., Fainsod A., Ruddle F.H.;
 RT "Sequence analysis of the murine Hox-2.2, -2.3, and -2.4 homeo boxes:
 evolutionary and structural comparisons.";
 RL Genomics 1:182-195(1987).
 RN [6]

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; X13961; CAA32141.1; -

DR EMBL; X13721; CAA32000.1; -

DR EMBL; M18399; ABA88246.1; ALT_SEQ.

DR EMBL; X54077; CAA38014.1; -

DR PIR; S03712; WJWS24.

DR HSP; P02833; 9ANT.

DR TRANSFAC; T01737; -

DR MGD; MGI:96189; Hoxb8.

DR InterPro; IPR001827; Antennapedia.

DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR000047; HTH_lamdrepressor.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00025; ANTENNAPEDIA.

DR PRINTS; PR00024; HOMEBOX.

DR PRINTS; PR00031; HTHREPRESSR.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; Hox; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS00071; HOMEBOX 2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;

KW Transcription regulation.

FT SITE 134 139 ANTP-TYPE HEXAPEPTIDE.

FT DNA_BIND 146 205 HOMEBOX.

SQ SEQUENCE 243 AA; 27618 MW; 368964C77884D21F CRC64;

Query Match 52.7%; Score 68; DB 1; Length 243;

Best Local Similarity 75.0%; Pred. No. 0.0046;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 18

HXA4_MORSA

ID HXA4_MORSA STANDARD; PRT; 248 AA.

QY 6 RQKIWFNRRKPKK 21
 |||||
 Db 188 RQKIWFNRRKPKK 203

RESULT 17

HXA4_HETFR

ID HXA4_HETFR STANDARD; PRT; 247 AA.

AC Q91A22;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Homeobox protein Hox-A4.

GN HXA4.

OS Heterodontus francisci (Horn shark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;

OC Heterodontidae; Heterodontus.

OX NCBI_TaxID=7792;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20144096; PubMed=10677514;

RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,

RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;

RT "Hox cluster genomics in the horn shark Heterodontus francisci.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis (By
 similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).

CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AF224262; AAF4642.1; -

DR HSP; P02833; 9ANT.

DR TRANSFAC; T0475; -

DR InterPro; IPR001827; Antennapedia.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00025; ANTENNAPEDIA.

DR PRINTS; PR00024; HOMEBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; Hox; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS00071; HOMEBOX 2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;

KW Transcription regulation.

FT SITE 128 133 ANTP-TYPE HEXAPEPTIDE.

FT DNA_BIND 149 208 HOMEBOX.

SQ SEQUENCE 247 AA; 28408 MW; BE518038CA73D2E6 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 247;

Best Local Similarity 75.0%; Pred. No. 0.0047;

Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQKIWFNRRKPKK 21
 |||||
 Db 191 RQKIWFNRRKPKK 206

RESULT 18

HXA4_MORSA

ID HXA4_MORSA STANDARD; PRT; 248 AA.

AC Q9PMD2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-A4.
 GN HoxA4.
 OS *Morone saxatilis* (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OC NCBI_TaxID=34816;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9259633; PubMed=10327649;
 RA Snell E.A., Scenama J.L., Stellwag E.J.;
 RT "Genomic organization of the Hoxa4-Hoxa10 region from *Morone saxatilis*: Implications for Hox gene evolution among vertebrates.";
 RL J. Exp. Zool. 285:41-49(1999).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of a developmental regulatory system that provides cells with specific positional identities on the anterior-posterior axis (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF089743; AAD46399.1; -;
 DR HSP; P02833; 9ANT.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEIDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEIDIA; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT SITE 128 133 ANTTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 149 208 HOMEBOX.
 SQ SEQUENCE 248 AA; 27729 MW; DAE182A92655F13 CRC64;
 Query Match 52.7%; Score 68; DB 1; Length 248;
 Best Local Similarity 75.0%; Pred. No. 0.0047;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 6 RPKIWFNRKPKWK 21
 DB 191 RQVKIWFQNRKPKWK 206
 RESULT 19
 HXA4 MOUSE
 ID HXA4 MOUSE STANDARD; PRT; 285 AA.
 AC P06798; Q61684; Q64388;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-A4 (Hox-1.4) (MH-3).
 GN HoxA4 OR HoxA-4 OR HOX-1.4.
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE OF 1-170 FROM N.A.
 RX MEDLINE=90214520; PubMed=2576648;
 RA Galliot B., Dolle P., Vigneron M., Featherstone M.S., Baron A., Duboule D.;
 RT "The mouse Hox-1.4 gene: primary structure, evidence for promoter activity and expression during development.";
 RL Development 107:343-359(1989).
 RN [2]
 RP SEQUENCE OF 171-285 FROM N.A.
 RX STRAIN=ICR; TISSUE=Spinal cord, and Testis;
 RA MEDLINE=92190549; PubMed=1686835;
 RA Rubin M.R., Nguyen-Huu M.C.;
 RT "Murine embryonic spinal cord and adult testis Hox-1.4 cDNAs are identical 3' to the homeo box.";
 RL DNA Seq. 1:329-334(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94227844; PubMed=7909702;
 RA Viviano C.M., Galliot B., Wolgemuth D.J.;
 RT "Multiple levels of regulation exist for expression of the Hoxa-4 (Hox-1.4) gene in the mouse testis.";
 RL Cell. Mol. Biol. Res. 39:483-495(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94043512; PubMed=7901228;
 RA Wu K., Wolgemuth D.J.;
 RT "Protein product of the somatic-type transcript of the Hoxa-4 (Hox-1.4) gene binds to homeobox consensus binding sites in its promoter and intron.";
 RL J. Cell. Biochem. 52:449-462(1993).
 RN [5]
 RP SEQUENCE OF 180-273 FROM N.A.
 RX MEDLINE=86261825; PubMed=3726554;
 RA Rubin M.R., Tott L.E., Patel M.D., D'Eustachio P., Nguyen-Huu M.C.;
 RT "A mouse homeo box gene is expressed in spermatocytes and embryos.";
 RL Science 233:663-667(1986).
 RN [6]
 RP SEQUENCE OF 198-271 FROM N.A.
 RX MEDLINE=86274625; PubMed=2426103;
 RA Wolgemuth D.J., Engelmyer E., Duggal R.N., Gizang-Ginsberg E., Mutter G.L., Ponzetto C., Viviano C., Zakari Z.F.;
 RT "Isolation of a mouse cDNA coding for a developmentally regulated, testis-specific transcript containing homeo box homology.";
 RL EMBO J. 5:1229-1235(1986).
 RN [7]
 RP SEQUENCE OF 180-285 FROM N.A.
 RX MEDLINE=88181884; PubMed=2895600;
 RA Duggal R.N., Zakari Z.F., Ponzetto C., Wolgemuth D.J.;
 RT "Differential expression of the c-abl proto-oncogene and the homeo box-containing gene Hox 1.4 during mouse spermatogenesis.";
 RL Ann. N.Y. Acad. Sci. 513:112-127(1987).
 RN [8]
 RP RESULTS OF OVEREXPRESSION.
 RX MEDLINE=89127494; PubMed=2563568;
 RA Wolgemuth D.J., Behringer R.R., Mostoller M.P., Brinster R.L., Palmiter R.D.;
 RT "Transgenic mice overexpressing the mouse homeobox-containing gene Hox-1.4 exhibit abnormal gut development.";
 RL Nature 337:464-467(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of a developmental regulatory system that provides cells with specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC HIGH LEVELS OF HOX-1.4 TRANSCRIPTS WERE FOUND AT DAYS 17, 24 AND 30. THE FIRST DAY OF HOX-1.4 EXPRESSION WAS DAY 14. THE ACTIVATION OF THE HOX-1.4 GENE IN MALE GERM CELLS SEEMS TO OCCUR AT THE PACHYTENE STAGE OF MEIOTIC PROPHASE AND ITS LEVEL OF EXPRESSION IS STAGE-SPECIFIC DURING EMBRYOGENESIS.
 CC -!- DISEASE: Overexpression results in abnormal gut development


```

DR SMART; SM00389; Hox; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 38 148 PRO-RICH.
FT SITE 188 193 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 209 268 HOMEBOX.
SQ SEQUENCE 309 AA; 33478 MW; 8898A41502319341 CRC64;

Query Match 52.7%; Score 68; DB 1; Length 309;
Best Local Similarity 75.0%; Pred. No. 0.0058;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RPKIWFNNRRKPKK 21
DB 251 RQVKIWFQRRMKKK 266

RESULT 23
HXAA4 HUMAN
ID HXAA4 HUMAN STANDARD; PRT; 320 AA.
AC Q00056; O43366;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-A4 (Hox-1D) (Hox-1.4).
GN HOXA4 OR HOXD4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91260707; PubMed=1675427;
RA Buettner R., Yim S.O., Hong Y.S., Boncinelli E., Tainsky M.A.;
RT "Alteration of homeobox gene expression by N-ras transformation of
PA-1 human teratocarcinoma cells.";
RL Mol. Cell. Biol. 11:3573-3583(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91153613; PubMed=1981366;
RA Peverali F.A., D'Esposito M., Acampora D., Bunone G., Negri M.,
RA Faiella A., Stornaiuolo A., Fammese M., Migliaccolo E., Simeone A.,
RA Valle G.D., Boncinelli E.;
RT "Expression of HOX homeogenes in human neuroblastoma cell culture
lines.";
RL Differentiation 45:61-69(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Bradshaw H., Hinds K., Keppler D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 215-280 FROM N.A.
RX MEDLINE=90215256; PubMed=2576652;
RA Boncinelli E., Acampora D., Fammese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RL Genome 31:745-756(1989).
RN [5]
RP SEQUENCE OF 205-277 FROM N.A.
RX MEDLINE=90007544; PubMed=2571574;
RA Ferguson-Smith A.C., Fienberg A., Ruddle F.H.;
RT "Isolation, chromosomal localization, and nucleotide sequence of the
human HOX 1.4 homeobox.";
RL Genomics 5:250-258(1989).
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC BINDS TO SITES IN THE 5'-FLANKING SEQUENCE OF ITS CODING REGION
CC WITH VARIOUS AFFINITIES. THE CONSENSUS SEQUENCES OF THE HIGH AND
CC LOW AFFINITY BINDING SITES ARE 5'-TAATGA[CG]-3' AND

```

```

CC 5'-CTAATTTT-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Embryonic nervous system.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74297; AAA58664.1; -.
DR EMBL; M28199; AAA53290.1; -.
DR EMBL; AC004080; -; NOT_ANNOTATED_CDS.
DR PIR; A39724; A39724.
DR HSSP; P02833; 9ANT.
DR TRANSFAC; T01703; -.
DR Genew; HGNC:5105; HOXA4.
DR MIM; 142953; -.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; Hox; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
DOMAIN 15 192 PRO-RICH; PART OF THE TRANSCRIPTIONAL
FT SITE 194 199 ACTIVATION DOMAIN.
FT DNA_BIND 215 274 ANTP-TYPE HEXAPEPTIDE.
FT CONFLICT 70 70 HOMEBOX.
FT CONFLICT 108 108 P -> T (IN REF. 3).
FT CONFLICT 140 140 A -> P (IN REF. 1).
FT CONFLICT 205 205 P -> L (IN REF. 3).
FT CONFLICT 275 275 A -> S (IN REF. 5).
FT CONFLICT 275 275 KL -> NC (IN REF. 5).
SQ SEQUENCE 320 AA; 34479 MW; A5E5C88108F0646D CRC64;

Query Match 52.7%; Score 68; DB 1; Length 320;
Best Local Similarity 75.0%; Pred. No. 0.006;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RPKIWFNNRRKPKK 21
DB 257 RQVKIWFQRRMKKK 272

RESULT 24
HXAA5 SHEEP
ID HXAA5 SHEEP STANDARD; PRT; 49 AA.
AC Q28599;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (Fragment).
GN HOXA5 OR HOXA-5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Roche P.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:14:23 ; Search time 36,3243 Seconds
(without alignments)
182.409 Million cell updates/sec

Title: US-09-977-349-4

Perfect score: 129

Sequence: 1 CSSCIRQPKIWFNRRKPMKK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPREMBL_25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_todent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	54.3	60	13	Q8QGL9
2	69	53.5	60	5	O77143
3	69	53.5	60	5	O77139
4	69	53.5	60	13	Q8QGL5
5	69	53.5	60	13	Q8QGL6
6	69	53.5	60	13	Q8QGL8
7	69	53.5	61	5	Q27910
8	69	53.5	77	5	O44260
9	69	53.5	89	5	Q967V4
10	69	53.5	94	5	Q967V6
11	69	53.5	107	5	Q967V8
12	69	53.5	115	5	O44258
13	69	53.5	115	13	Q801B5
14	69	53.5	146	5	Q05009
15	69	53.5	244	13	Q801C1
16	69	53.5	264	4	Q86TF7

17	68	52.7	39	6	O9XT68	O9xt68 canis famil
18	68	52.7	44	5	O9GR21	O9gr21 nematostell
19	68	52.7	60	5	Q26375	Q26375 styela clav
20	68	52.7	60	11	Q80WH3	Q80wh3 rattus sp.
21	68	52.7	60	11	Q80WH2	Q80wh2 rattus sp.
22	68	52.7	60	13	O9PSZ1	O9ps21 petromyzon
23	68	52.7	60	13	O91769	O91769 xenopus lae
24	68	52.7	60	13	O9PSZ2	O9ps22 petromyzon
25	68	52.7	63	5	O8WRM1	O8wrml lithobius a
26	68	52.7	69	13	O9PVR3	O9pvr3 oryzias lat
27	68	52.7	74	13	O57369	O57369 brachydanio
28	68	52.7	81	13	O9PVS5	O9pvs5 oryzias lat
29	68	52.7	82	13	O9PVR4	O9pvr4 oryzias lat
30	68	52.7	84	13	O57363	O57363 brachydanio
31	68	52.7	85	5	O9GP50	O9gps0 discocelis
32	68	52.7	85	11	O63255	O63255 rattus norv
33	68	52.7	85	11	O61681	O61681 mus musculu
34	68	52.7	86	13	O57360	O57360 brachydanio
35	68	52.7	89	13	O42371	O42371 brachydanio
36	68	52.7	90	5	O817C9	O817c9 ciona intes
37	68	52.7	93	13	O801B3	O801b3 latimeria m
38	68	52.7	103	13	O93286	O93286 ambystoma m
39	68	52.7	104	13	O9PMM1	O9pwm1 brachydanio
40	68	52.7	106	13	O801D2	O801d2 latimeria m
41	68	52.7	125	5	O96896	O96896 sacculina c
42	68	52.7	141	11	O8BQA3	O8bqa3 mus musculu
43	68	52.7	155	5	O25577	O25577 nematostell
44	68	52.7	176	11	O8BNV9	O8bnv9 mus musculu
45	68	52.7	181	5	Q26611	Q26611 styela plic
46	68	52.7	182	5	O9U8Q5	O9ubq5 dicyema ori
47	68	52.7	185	11	O8BNX6	O8bnx6 mus musculu
48	68	52.7	193	5	O817D0	O817d0 ciona intes
49	68	52.7	194	5	O26478	O26478 styela clav
50	68	52.7	218	11	O9CUH8	O9cuh8 mus musculu
51	68	52.7	241	13	O9YH27	O9yh27 gallus gall
52	68	52.7	242	13	O9YH13	O9yh13 gallus gall
53	68	52.7	243	4	O8N8T3	O8n8t3 homo sapien
54	68	52.7	244	13	O9PWL5	O9pwl5 brachydanio
55	68	52.7	245	13	O8AWZ0	O8awz0 brachydanio
56	68	52.7	247	13	O8JH55	O8jhs5 brachydanio
57	68	52.7	272	13	O9PSW4	O9psw4 petromyzon
58	68	52.7	285	11	O8BPE6	O8bpe6 mus musculu
59	68	52.7	289	4	O81XZ1	O8ixz1 homo sapien
60	68	52.7	363	5	O9XY00	O9xy00 dugesia jap
61	68	52.7	392	13	O42503	O42503 figu rubrip
62	68	52.7	435	5	O817C8	O817c8 ciona intes
63	67	51.9	33	5	O86FU0	O86fu0 grosophila
64	67	51.9	39	13	O57368	O57368 brachydanio
65	67	51.9	42	11	O80WH6	O80wh6 rattus sp.
66	67	51.9	43	13	O57359	O57359 brachydanio
67	67	51.9	46	13	O9PVR9	O9pvr9 oryzias lat
68	67	51.9	51	5	O27413	O27413 ctenodrilus
69	67	51.9	51	5	O23743	O23743 ctenodrilus
70	67	51.9	51	5	O26407	O26407 ctenodrilus
71	67	51.9	57	13	O9PVR8	O9pvr8 oryzias lat
72	67	51.9	58	5	O9Y188	O9y188 priapulua c
73	67	51.9	58	5	O25208	O25208 junonia coe
74	67	51.9	58	13	O57362	O57362 brachydanio
75	67	51.9	59	5	O8WRM9	O8wrms lithobius a
76	67	51.9	59	5	O9NR42	O9nrb2 anopheles g
77	67	51.9	59	13	O9PVR5	O9pvr5 oryzias lat
78	67	51.9	60	11	O80WH7	O80wh7 rattus sp.
79	67	51.9	60	13	O80WH4	O80wh4 rattus sp.
80	67	51.9	60	13	O8QGL3	O8qgl3 petromyzon
81	67	51.9	60	13	O8QGL2	O8qgl2 petromyzon
82	67	51.9	60	13	O8QGL7	O8qgl7 petromyzon
83	67	51.9	63	5	O77138	O77138 archegozete
84	67	51.9	63	5	O8MXB2	O8mxb2 holopneuste
85	67	51.9	66	13	O57356	O57356 brachydanio
86	67	51.9	69	5	O9U9T4	O9u9t4 hereis wire
87	67	51.9	69	5	O9BMF7	O9bmf7 hallois as
88	67	51.9	70	5	O967W5	O967w5 folsonia ca
89	67	51.9	70	13	Q801B4	Q801b4 latimeria m

DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7755 MW; 32678A250BDEF74 CRC64;
 Query Match
 Best Local Similarity 53.5%; Score 69; DB 5; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58
 RESULT 4
 Q8QGL5
 ID Q8QGL5 PRELIMINARY; PRT; 60 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hox15/6 homeobox (Fragment).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 marinus.";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF410913; AAM19471.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7715 MW; 17A6ED79EF76A4F5 CRC64;
 Query Match
 Best Local Similarity 53.5%; Score 69; DB 13; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58
 RESULT 5
 Q8QGL6
 ID Q8QGL6 PRELIMINARY; PRT; 60 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hox16 homeobox (Fragment).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 marinus.";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF410914; AAM19472.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7717 MW; E150F172FD751F3C CRC64;
 Query Match
 Best Local Similarity 53.5%; Score 69; DB 13; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58
 RESULT 6
 Q8QGL8
 ID Q8QGL8 PRELIMINARY; PRT; 60 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE HoxG4 homeobox (Fragment).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 marinus.";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF410911; AAM19469.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7715 MW; 17A6ED79EF76A4F5 CRC64;
 Query Match
 Best Local Similarity 53.6%; Score 69; DB 13; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 marinus.";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF410913; AAM19471.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7715 MW; 17A6ED79EF76A4F5 CRC64;
 Query Match
 Best Local Similarity 53.5%; Score 69; DB 13; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58
 RESULT 6
 Q8QGL8
 ID Q8QGL8 PRELIMINARY; PRT; 60 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE HoxG4 homeobox (Fragment).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 marinus.";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF410911; AAM19469.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepressr.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON TER 1 1
 FT NON TER 60 60
 SQ SEQUENCE 60 AA; 7715 MW; 17A6ED79EF76A4F5 CRC64;
 Query Match
 Best Local Similarity 53.6%; Score 69; DB 13; Length 60;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SSC1--RQPKIWFNNRRKPKWK 21
 Db 37 SLCLTERQIKIWFQNNRMKWK 58


```

FT NON TER 60 60
SQ SEQUENCE 60 AA; 7704 MW; 0EBB642C24DEC6 CRC64;

Query Match 53.5%; Score 69; DB 13; Length 60;
Best Local Similarity 63.6%; Pred. NO. 0.0062;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKK 21
   | : | | | | | | | | | |
Db 37 SLCLSERQIKIWFQNRKMKKK 58

RESULT 7
Q27910
ID Q27910 PRELIMINARY; PRT; 61 AA.
AC Q27910; Q94684;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeodomain protein PHOX5 (Fragment).
GN PHOX5.
OS Polyandrocampa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocampa.
OX NCB1_TaxID=7723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USHIMADO;
RA Fujiwara S., Kawamura K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-25 FROM N.A.
RC STRAIN=USHIMADO;
RA Fujiwara S., Kawamura K.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; D86851; BAAL3173.1; -.
DR EMBL; D88277; BAAL3573.1; -.
DR HSSP; P02833; LHOM.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambda repressor.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 61 AA; 7537 MW; 14C09C6BAF112C33 CRC64;

Query Match 53.5%; Score 69; DB 5; Length 61;
Best Local Similarity 63.6%; Pred. NO. 0.0063;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKK 21
   | : | | | | | | | | | |
Db 17 SLCLSERQIKIWFQNRKMKKK 38

RESULT 8
Q44260
ID Q44260 PRELIMINARY; PRT; 77 AA.
AC Q44260;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sex combs reduced (Fragment).
GN SCR.

```

```

OS Ethmostigmus rubripes.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Ethmostigmus.
OX NCB1_TaxID=62613;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97411147; PubMed=9259556;
RA Grenier J.K., Garber T.L., Warren R., Whittington P.M., Carroll S.;
RT "Evolution of the entire arthropod Hox gene set predated the origin
RT and radiation of the onychophoran/arthropod clade.";
RL Curr. Biol. 7:547-553(1997).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF010178; AAB91392.1; -.
DR HSSP; P02833; IAHD.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 77 AA; 9377 MW; DF13885AD5B5EB71 CRC64;

Query Match 53.5%; Score 69; DB 5; Length 77;
Best Local Similarity 63.6%; Pred. NO. 0.0078;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKK 21
   | : | | | | | | | | | |
Db 31 SLCLSERQIKIWFQNRKMKKK 52

RESULT 9
Q967V4
ID Q967V4 PRELIMINARY; PRT; 89 AA.
AC Q967V4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sex combs reduced 2 (Fragment).
GN SCR2.
OS Lithobius forficatus.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCB1_TaxID=7552;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21272202; PubMed=11378385;
RA Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M.E.;
RT "Hox genes and phylogeny of the arthropods.";
RL Curr. Biol. 11:759-763(2001).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF362089; AAK51944.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODOM; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 89 AA; 10703 MW; 240F5DAA5307AC2C CRC64;

Query Match 53.5%; Score 69; DB 5; Length 89;

```

```

Best Local Similarity 63.6%; Pred. No. 0.009; Mismatches 1; Indels 5; Gaps 1;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKWK 21
Db 43 SLCLSERQIKIWFNNRRKPKWK 64

RESULT 10
Q967V6 PRELIMINARY; PRT; 94 AA.
AC Q967V6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deformed (Fragment).
GN DFD.
OS Lithobius forficatus.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=7552;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=2127202; PubMed=11378385;
RA Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M.E.;
RT "Hox genes and phylogeny of the arthropods."
RL Curr. Biol. 11:758-763 (2001).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF362087; AAKS1942.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 11485 MW; 67DFFE7DAF3B850F CRC64;

Query Match 53.5%; Score 69; DB 5; Length 94;
Best Local Similarity 63.6%; Pred. No. 0.0094;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKWK 21
Db 46 SLCLSERQIKIWFNNRRKPKWK 67

RESULT 11
Q967V6 PRELIMINARY; PRT; 107 AA.
AC Q967V6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deformed (Fragment).
GN DFD.
OS Lithobius atkinsoni.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
OX NCBI_TaxID=177213;
RN [1]_
RP SEQUENCE FROM N.A.
RX Hughes C.L., Kaufman T.C.;
RA "Exploring the myriapod body plan: expression patterns of the ten Hox
RT genes in a centipede."
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

```

```

DR EMBL; AF434997; AAL36902.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 13030 MW; 895A860B7E904DBD CRC64;

Query Match 53.5%; Score 69; DB 5; Length 107;
Best Local Similarity 63.6%; Pred. No. 0.011;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKWK 21
Db 28 SLCLSERQIKIWFNNRRKPKWK 49

RESULT 12
O44258 PRELIMINARY; PRT; 115 AA.
AC O44258;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Deformed (Fragment).
GN DFD.
OS Ethmostigmus rubripes.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Ethmostigmus.
OX NCBI_TaxID=62613;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=97411147; PubMed=9259556;
RA Grenier J.K., Garber T.L., Warren R., Whittington P.M., Carroll S.;
RT "Evolution of the entire arthropod Hox gene set predated the origin
RT and radiation of the onychophoran/arthropod clade."
RL Curr. Biol. 7:547-553 (1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF010176; AAB91390.1; -.
DR HSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 13633 MW; DOAECL369ACF517 CRC64;

Query Match 53.5%; Score 69; DB 5; Length 115;
Best Local Similarity 63.6%; Pred. No. 0.011;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNNRRKPKWK 21
Db 17 SLCLSERQIKIWFNNRRKPKWK 38

RESULT 13
Q801B5 PRELIMINARY; PRT; 115 AA.
ID Q801B5

```

```

AC Q801B5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HoxC4 (Fragment)
OS Latimeria menadoensis (Indonesian coelacanth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OX NCBI_TaxID=106881;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22457206; PubMed=12547909;
RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RT "Hox gene clusters in the Indonesian coelacanth, Latimeria
RT menadoensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1084-1088 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Koh E.G.L., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AY183742; AAO43035.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD00024; HOMEBOX.
DR PRINTS; PS00027; HOMEBOX_1; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 13462 MW; 956057D9822E9B52 CRC64;

Query Match 53.5%; Score 69; DB 13; Length 115;
Best Local Similarity 63.6%; Pred. No. 0.011;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNRRKPKWK 21
Db | : | | | | | | | | | | | | | | | | | | | |
45 SLCLSERQIKIWFQNRKMKWK 66

RESULT 14
ID Q05009 PRELIMINARY; PRT; 146 AA.
AC Q05009;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dfd protein (Fragment).
GN DFD.
OS Artemia salinifranciscana (Brine shrimp) (Artemia franciscana).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OX NCBI_TaxID=6661;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=North arm of Great Salt Lake;
RA Averof M., Akam M.;
RT "HOM/HOX genes in a crustacean: implication for the origin of insect
RT and crustacean body plans.";
RL Curr. Biol. 3:73-78 (1993).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; X70078; CAA49683.1; -.
DR HSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.

```

```

DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 146 AA; 16972 MW; 03D47AD09575C129 CRC64;

Query Match 53.5%; Score 69; DB 5; Length 146;
Best Local Similarity 63.6%; Pred. No. 0.014;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNRRKPKWK 21
Db | : | | | | | | | | | | | | | | | | | | | |
47 SLCLSERQIKIWFQNRKMKWK 68

RESULT 15
ID Q801C1 PRELIMINARY; PRT; 244 AA.
AC Q801C1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HoxB4.
OS Latimeria menadoensis (Indonesian coelacanth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OX NCBI_TaxID=106881;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22457206; PubMed=12547909;
RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RT "Hox gene clusters in the Indonesian coelacanth, Latimeria
RT menadoensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1084-1088 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Koh E.G.L., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AY183736; AAO43029.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
SQ SEQUENCE 244 AA; 28024 MW; E761355D928226D6 CRC64;

Query Match 53.5%; Score 69; DB 13; Length 244;
Best Local Similarity 63.6%; Pred. No. 0.023;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNRRKPKWK 21
Db | : | | | | | | | | | | | | | | | | | | | |
183 SLCLSERQIKIWFQNRKMKWK 204

RESULT 16
ID Q86TF7 PRELIMINARY; PRT; 264 AA.
AC Q86TF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

RA	SEQUENCE FROM N.A.	
RP	MEDLINE=99265967; PubMed=10331940;	
RX	Li R., Mignot E., Faraco J., Kadotani H., Cantanese J., Zhao B.,	
RA	Lin X., Hinton L., Osterander E.A., Patterson D.F., de Jong P.J.;	
RA	"Construction and characterization of an eightfold redundant dog	
RT	genomic bacterial artificial chromosome library."	
RL	Genomics 58:9-17(1999).	
RN	[2]	
RA	SEQUENCE FROM N.A.	
RA	Kodatani H., Mignot E.;	
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	
CC	- - SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).	
DR	EMBL; AF103746; AAD40572.1; -	
DR	GO; GO:0005634; C:nucleus; IEA.	
DR	GO; GO:0003700; F:transcription factor activity; IEA.	
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001356; Homeobox.	
DR	Pfam; PF00046; homeobox; 1.	
DR	PRINTS; PS00024; HOMEBOX.	
DR	ProDom; PD00010; Homeobox; 1.	
DR	SMART; SM00389; HOX; 1.	
DR	PROSITE; PS00027; HOMEBOX 1; 1.	
DR	PROSITE; PS00071; HOMEBOX 2; 1.	
KW	DNA-binding; Homeobox; Nucleic protein.	
FT	NON_TER 1 1	
FT	NON_TER 39 39	
FT	SEQUENCE 39 AA; 4908 MW; 1252D75EA245BCAB CRC64;	
SQ	Query Match 52.7%; Score 68; DB 6; Length 39;	
	Best Local Similarity 75.0%; Pred. No. 0.0056;	
	Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps	
Oy	6 RQKWIWFPNRRKPWK 21	
Db	13 RQKWIWFPNRRKPWK 28	
RESULT 18		
Q9GR21	PRELIMINARY; PRT; 44 AA.	
ID	O9GR21	
AC	O9GR21	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
DE	Hox type homeodomain protein (Fragment).	
GN	ANTHOX1A.	
OS	Nematostella vectensis.	
OC	Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;	
OC	Edwardsiidae; Nematostella.	
OX	NCBI_TaxID=45351;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=97435515; PubMed=9290214;	
RA	Finnerty J.R., Martindale M.Q.;	
RA	"Homeoboxes in sea anemones (Chnidaria:Anthozoa): a PCR-based survey of	
RT	Nematostella vectensis and Metridium senile.;"	
RL	Biol. Bull. 193:62-76(1997).	
CC	- - SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).	
DR	EMBL; AF020955; AAG37789.1; -	
DR	GO; GO:0005634; C:nucleus; IEA.	
DR	GO; GO:0003700; F:transcription factor activity; IEA.	
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001356; Homeobox.	
DR	Pfam; PF00046; homeobox; 1.	
DR	PRINTS; PR00024; HOMEBOX.	
DR	ProDom; PD00010; Homeobox; 1.	
DR	SMART; SM00389; HOX; 1.	
DR	PROSITE; PS00027; HOMEBOX 1; 1.	
DR	PROSITE; PS00071; HOMEBOX 2; 1.	
KW	DNA-binding; Homeobox; Nucleic protein.	
FT	NON_TER 1 1	
FT	SEQUENCE 44 AA; 5362 MW; PAF8044070A9CFC3A CRC64.	


```

Db      43 RQVKWIFQNRKMKWK 58
      || ||| ||| ||| |||
RESULT 22
Q9PSZ1 ID Q9PSZ1 PRELIMINARY; PRT; 60 AA.
AC Q9PSZ1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox protein HoxQa (Fragment).
GN HOXQ8A.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98100508; PubMed=9437854;
RA Carr J.L., Shashikant C.S., Bailey W.J., Ruddle F.H.;
RT "Molecular evolution of Hox gene regulation: cloning and transgenic
analysis of the lamprey Hox8 gene.";
RL J. Exp. Zool. 280:73-85 (1998).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF035589; AAC04331.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR IPR000047; HTH lambdarepressr.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7655 MW; DD30A4DBA91F1F58 CRC64;

Query Match 52.7%; Score 68; DB 13; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKWFPNRRKPKWK 21
Db 43 RQVKWIFQNRKMKWK 58
      || ||| ||| ||| |||
RESULT 23
Q91769 ID Q91769 PRELIMINARY; PRT; 60 AA.
AC Q91769
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeodomain protein (Fragment).
GN HOMEBOX.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93326800; PubMed=7916675;
RA Bittner D.A., De Robertis E.W., Cho K.W.;
RT "Characterization of the Xenopus Hox 2.4 gene and identification of
control elements in its intron.";

```

```

RL Dev. Dyn. 136:11-24 (1993).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; M83947; AAA49748.1; -.
DR PIR; I51437; I51437.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR IPR000047; HTH lambdarepressr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7671 MW; DD27AFB608CF1F58 CRC64;

Query Match 52.7%; Score 68; DB 13; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKWFPNRRKPKWK 21
Db 43 RQVKWIFQNRKMKWK 58
      || ||| ||| ||| |||
RESULT 24
Q9PSZ2 ID Q9PSZ2 PRELIMINARY; PRT; 60 AA.
AC Q9PSZ2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox protein HoxR8 (Fragment).
GN HOXR8.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98100508; PubMed=9437854;
RA Carr J.L., Shashikant C.S., Bailey W.J., Ruddle F.H.;
RT "Molecular evolution of Hox gene regulation: cloning and transgenic
analysis of the lamprey Hox8 gene.";
RL J. Exp. Zool. 280:73-85 (1998).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF035588; AAC04330.1; -.
DR HSSP; P02833; 2HOA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR IPR000047; HTH lambdarepressr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR PRODom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 7685 MW; 2C40AFB3089F1F40 CRC64;

Query Match 52.7%; Score 68; DB 13; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Tue May 25 09:20:21 2004

QY 6 RQPKIWFNRRKPWK 21
 |||||
 Db 43 RQVKIWFQNRMRWK 58
 |||||

RESULT 25

Q8WRM1 PRELIMINARY; PRT; 63 AA.
 AC Q8WRM1;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Sex combs reduced (Fragment).
 GN SCR.
 OS Lithobius atkinsoni.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
 OX NCBI_TaxID=17213;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hughes C.L., Kaufman T.C.;
 RT "Exploring the myriapod body plan: expression patterns of the ten Hox
 genes in a centipede.";
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF435004; AAL36909.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS0071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON TER 1
 SQ SEQUENCE 63 AA; 7616 MW; 34FFC51F793427FD CRC64;

Query Match 52.7%; Score 68; DB 5; Length 63;
 Best Local Similarity 63.6%; Pred. No. 0.0091;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKIWFNRRKPWK 21
 |||||
 Db 17 SLCLSERQIKIWFQNRMRWK 38
 |||||

Search completed: May 24, 2004, 17:24:41
 Job time : 37.3243 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:22:27 ; Search time 15.0405 Seconds
(without alignments)
72.082 Million cell updates/sec

Title: US-09-977-349-4

Perfect score: 129

Sequence: 1 CSSCIRQPKIWFPPNRKPKWK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*

2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*

3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*

4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*

5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep.*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	76.7	16	3	US-08-849-486-6
2	75	58.1	16	3	US-08-849-486-5
3	70	54.3	22	4	US-09-155-165-11
4	68	52.7	61	3	US-08-751-344B-7
5	67	51.9	16	2	US-08-928-958-7
6	67	51.9	16	2	US-08-810-540-3
7	67	51.9	16	2	US-08-810-540-6
8	67	51.9	16	2	US-09-072-429-7
9	67	51.9	16	3	US-08-964-302A-6
10	67	51.9	16	3	US-09-116-294-4
11	67	51.9	16	3	US-08-964-614A-4
12	67	51.9	16	3	US-08-849-486-1
13	67	51.9	16	3	US-08-849-486-4
14	67	51.9	16	3	US-09-208-966-54
15	67	51.9	16	3	US-09-308-935-8
16	67	51.9	16	3	US-09-441-416A-6
17	67	51.9	16	4	US-09-296-089-33
18	67	51.9	16	4	US-09-419-826-35
19	67	51.9	16	4	US-09-302-305C-10
20	67	51.9	16	4	US-09-346-847-1
21	67	51.9	16	4	US-09-346-847-25
22	67	51.9	16	4	US-09-057-363C-47
23	67	51.9	16	4	US-09-043-560B-3
24	67	51.9	16	4	US-09-648-400A-29
25	67	51.9	16	4	US-09-227-652B-4
26	67	51.9	16	4	US-09-780-070-38
27	67	51.9	16	4	US-08-610-220B-9

28	67	51.9	16	4	US-09-775-052A-54	Sequence 54, Appl
29	67	51.9	16	4	US-09-155-165-22	Sequence 22, Appl
30	67	51.9	16	4	US-09-792-480-29	Sequence 29, Appl
31	67	51.9	16	4	US-09-551-976-33	Sequence 33, Appl
32	67	51.9	16	4	US-09-265-107-47	Sequence 47, Appl
33	67	51.9	17	4	US-09-346-847-17	Sequence 17, Appl
34	67	51.9	17	4	US-09-346-847-20	Sequence 20, Appl
35	67	51.9	17	4	US-09-346-847-22	Sequence 22, Appl
36	67	51.9	17	4	US-09-346-847-27	Sequence 27, Appl
37	67	51.9	17	4	US-09-648-400A-30	Sequence 30, Appl
38	67	51.9	18	3	US-08-838-545-20	Sequence 20, Appl
39	67	51.9	18	3	US-09-349-532-20	Sequence 20, Appl
40	67	51.9	19	4	US-09-346-847-23	Sequence 23, Appl
41	67	51.9	19	4	US-09-658-517C-7	Sequence 7, Appl
42	67	51.9	20	4	US-09-346-847-16	Sequence 16, Appl
43	67	51.9	20	4	US-09-346-847-18	Sequence 18, Appl
44	67	51.9	20	4	US-09-346-847-30	Sequence 30, Appl
45	67	51.9	20	4	US-09-658-517C-8	Sequence 8, Appl
46	67	51.9	21	4	US-08-610-220B-11	Sequence 11, Appl
47	67	51.9	22	4	US-09-346-847-28	Sequence 28, Appl
48	67	51.9	22	4	US-09-057-363C-50	Sequence 50, Appl
49	67	51.9	22	4	US-08-610-220B-10	Sequence 10, Appl
50	67	51.9	22	4	US-09-155-165-5	Sequence 5, Appl
51	67	51.9	22	4	US-09-265-107-50	Sequence 50, Appl
52	67	51.9	24	4	US-09-419-826-34	Sequence 34, Appl
53	67	51.9	24	4	US-09-428-082B-332	Sequence 332, Appl
54	67	51.9	27	3	US-09-051-934-51	Sequence 51, Appl
55	67	51.9	27	3	US-09-051-934-52	Sequence 52, Appl
56	67	51.9	27	4	US-09-040-725A-2	Sequence 2, Appl
57	67	51.9	34	4	US-09-347-504-79	Sequence 79, Appl
58	67	51.9	34	4	US-10-161-499-79	Sequence 79, Appl
59	67	51.9	36	4	US-09-428-082B-331	Sequence 331, Appl
60	67	51.9	61	2	US-08-202-044-3	Sequence 3, Appl
61	67	51.9	61	3	US-08-751-344B-3	Sequence 3, Appl
62	67	51.9	61	3	US-08-751-344B-6	Sequence 6, Appl
63	67	51.9	61	3	US-08-751-344B-9	Sequence 9, Appl
64	66	51.2	16	4	US-09-466-772-7	Sequence 7, Appl
65	66	51.2	20	4	US-09-466-772-3	Sequence 3, Appl
66	66	51.2	21	4	US-09-466-772-1	Sequence 1, Appl
67	66	51.2	22	4	US-09-466-772-2	Sequence 2, Appl
68	66	51.2	23	4	US-09-466-772-4	Sequence 4, Appl
69	64	49.6	16	3	US-08-849-486-7	Sequence 7, Appl
70	64	49.6	302	3	US-08-203-532F-4	Sequence 4, Appl
71	64	49.6	302	3	US-08-950-860-16	Sequence 16, Appl
72	64	49.6	302	3	US-09-078-465-4	Sequence 4, Appl
73	64	49.6	302	5	PCT-US95-01882A-4	Sequence 4, Appl
74	64	49.6	303	3	US-08-203-532F-2	Sequence 2, Appl
75	64	49.6	303	3	US-09-078-465-2	Sequence 2, Appl
76	64	49.6	303	5	PCT-US95-01882A-2	Sequence 2, Appl
77	62	48.1	15	2	US-08-810-540-4	Sequence 4, Appl
78	62	48.1	16	4	US-09-792-480-30	Sequence 30, Appl
79	62	48.1	42	3	US-08-751-344B-4	Sequence 4, Appl
80	62	48.1	283	1	US-08-583-672-2	Sequence 2, Appl
81	62	48.1	283	2	US-08-202-044-2	Sequence 2, Appl
82	62	48.1	283	3	US-08-751-344B-2	Sequence 2, Appl
83	62	48.1	284	3	US-08-320-148B-2	Sequence 2, Appl
84	62	48.1	284	3	US-08-589-028-6	Sequence 6, Appl
85	62	48.1	284	3	US-08-784-582-6	Sequence 6, Appl
86	62	48.1	284	3	US-08-785-271-6	Sequence 6, Appl
87	62	48.1	284	3	US-09-031-898-2	Sequence 2, Appl
88	60	46.5	16	3	US-08-849-486-8	Sequence 8, Appl
89	60	46.5	16	3	US-09-208-966-1	Sequence 1, Appl
90	60	46.5	16	4	US-09-775-052A-1	Sequence 1, Appl
91	60	46.5	300	4	US-09-162-524-1	Sequence 1, Appl
92	60	46.5	349	4	US-09-162-524-3	Sequence 3, Appl
93	59	45.7	367	4	US-09-009-816-2	Sequence 2, Appl
94	58	45.0	16	4	US-09-296-083-34	Sequence 34, Appl
95	58	45.0	16	4	US-09-057-363C-48	Sequence 48, Appl
96	58	45.0	16	4	US-09-551-976-34	Sequence 34, Appl
97	58	45.0	16	4	US-09-265-107-48	Sequence 48, Appl
98	58	45.0	22	4	US-09-265-107-72	Sequence 72, Appl
99	57	44.2	16	4	US-09-402-929-7	Sequence 7, Appl
100	56	43.4	15	4	US-10-083-889-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
 US-08-849-486-6
 ; Sequence 6, Application US/08849486
 ; Patent No. 6080724
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
 ; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
 ; NUMBER OF SEQUENCES: 10
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/849,486
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 95 11714
 ; FILING DATE: 05-OCT-1995
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-849-486-6

Query Match 76.7%; Score 99; DB 3; Length 16;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 RQPKIWFPPNRKPKWK 21
 DB 1 RQPKIWFPPNRKPKWK 16
 |||||

RESULT 2
 US-08-849-486-5
 ; Sequence 5, Application US/08849486
 ; Patent No. 6080724
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
 ; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
 ; NUMBER OF SEQUENCES: 10
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/849,486
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 95 11714
 ; FILING DATE: 05-OCT-1995
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-849-486-5

Query Match 58.1%; Score 75; DB 3; Length 16;
 Best Local Similarity 81.2%; Pred. No. 0.00012;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 6 RQPKIWFPPNRKPKWK 21
 DB 1 RQPKIWFPPNRKPKWK 16
 |||||

RESULT 3
 US-09-155-165-11
 ; Sequence 11, Application US/09155165
 ; Patent No. 6660830
 ; GENERAL INFORMATION:
 ; APPLICANT: Radulescu, Razvan T
 ; TITLE OF INVENTION: PEPTIDES WITH ANTIPROLIFERATIVE PROPERTIES
 ; FILE REFERENCE: 201196/20
 ; CURRENT APPLICATION NUMBER: US/09/155,165
 ; CURRENT FILING DATE: 1999-06-07
 ; PRIOR APPLICATION NUMBER: 09/155,165
 ; PRIOR FILING DATE: 1998-09-22
 ; PRIOR APPLICATION NUMBER: PCT/DE97/00643
 ; PRIOR FILING DATE: 1997-03-26
 ; PRIOR APPLICATION NUMBER: DE 196 11 939.1
 ; PRIOR FILING DATE: 1996-03-26
 ; PRIOR APPLICATION NUMBER: DE 196 53 445.3
 ; PRIOR FILING DATE: 1996-12-20
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 11
 ; LENGTH: 22
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Peptide
 ; NAME/KEY: UNSURE
 ; LOCATION: (1)..(22)
 ; OTHER INFORMATION: Where all amino acids may be in L or D
 ; OTHER INFORMATION: configuration
 US-09-155-165-11

Query Match 54.3%; Score 70; DB 4; Length 22;
 Best Local Similarity 70.6%; Pred. No. 0.00088;
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 5 IROPKIWFPPNRKPKWK 21
 DB 6 VRQIKIWFQNRKPKWK 22
 :|||

RESULT 4
 US-08-751-344B-7
 ; Sequence 7, Application US/08751344B
 ; Patent No. 6210960
 ; GENERAL INFORMATION:
 ; APPLICANT: Habener M.D., Joel F.
 ; APPLICANT: Miller Ph.D., Christopher P.
 ; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Banner & Witcoff, Ltd.
 ; STREET: One Financial Center
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WordPerfect 6.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/751,344B
;; FILING DATE: 19-NOV-2003
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/202,044
;; FILING DATE: 23-FEB-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Williams Ph.D., Kathleen M.
;; REGISTRATION NUMBER: 34,380
;; REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
;; TELEPHONE: (617) 345-9100
;; TELEFAX: (617) 345-9111
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 61 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
US-08-751-344B-7

Query Match 52.7%; Score 68; DB 3; Length 61;
Best Local Similarity 75.0%; Pred. No. 0.0049; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 4;

QY 6 RQPKIWFNRRKPKWK 21
DB 44 RQKIFQNRKMKWK 59

RESULT 5
US-08-928-958-7
; Sequence 7, Application US/08928958
; Patent No. 5877282
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CLEAVELAND, JEFFREY S.
; APPLICANT: BLAKE, JAMES
; APPLICANT: HAFKAR, OMAR K.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,958
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026978
; FILING DATE: 20-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 7:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-928-958-7

Query Match 51.9%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
DB 1 RQKIFQNRKMKWK 16

RESULT 6
US-08-810-540-3
; Sequence 3, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
; APPLICANT: TROY, CAROL M.
; APPLICANT: Shelanski, Michael L.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,540
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-810-540-3

Query Match 51.9%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
DB 1 RQKIFQNRKMKWK 16

RESULT 7
US-08-810-540-6
; Sequence 6, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:

APPLICANT: Troy, Carol M.
APPLICANT: Shelanski, Michael L.
TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
TITLE OF INVENTION: DEATH AND USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,540
FILING DATE: 03-MAR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/51247
TELEPHONE: 212-378-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-810-540-6

Query Match 51.9%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKWFNNRRKPKWK 21
Db 1 RQKWFNNRRKPKWK 16

RESULT 8
US-09-072-429-7
Sequence 7, Application US/09072429
Patent No. 5962415
GENERAL INFORMATION:
APPLICANT: Nadler, Steven G.
TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: P. O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,429
FILING DATE: 04-MAY-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Klein, Christopher A.
REGISTRATION NUMBER: 34,363
REFERENCE/DOCKET NUMBER: ON0141b
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 252-3714
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-072-429-7

Query Match 51.9%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKWFNNRRKPKWK 21
Db 1 RQKWFNNRRKPKWK 16

RESULT 9
US-08-964-302A-6
Sequence 6, Application US/08964302A
Patent No. 6015787
GENERAL INFORMATION:
APPLICANT: Potter, David A.
APPLICANT: Skolnik, Paul R.
TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,302A
FILING DATE: 04-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 00398/126001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-964-302A-6

Query Match 51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKWFNNRRKPKWK 21
Db 1 RQKWFNNRRKPKWK 16

```
RESULT 10
US-09-116-294-4
; Sequence 4, Application US/09116294
; Patent No. 6025140
; GENERAL INFORMATION:
; APPLICANT: Langel, Ulo
; APPLICANT: Bartfai, Tamás
; APPLICANT: Pooga, Margus
; APPLICANT: Valkna, Andres
; APPLICANT: Saar, Kulliki
; APPLICANT: Hallbrink, Mattias
; TITLE OF INVENTION: Conjugated Constructs of Peptides and
; TITLE OF INVENTION: Nucleic Acid Analogs, and Their Transport Across Membranes
; FILE REFERENCE: 4394
; CURRENT APPLICATION NUMBER: US/09/116,294
; CURRENT FILING DATE: 1998-07-16
; EARLIER APPLICATION NUMBER: 60/052,678
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: drosophila
US-09-116-294-4

Query Match 51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
Db 1 RQIKIWFQNRMRKWK 16

RESULT 11
US-08-964-614A-4
; Sequence 4, Application US/08964614A
; Patent No. 6057104
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN
; TITLE OF INVENTION: RAD51 PROTEIN AND DISRUPTION OF PROTEINS THAT ASSOCIATE
; TITLE OF INVENTION: WITH MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,614A
; FILING DATE: 05-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,280
; FILING DATE: 05-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8535-0019-999
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:

Query Match 51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
Db 1 RQIKIWFQNRMRKWK 16

RESULT 12
US-08-849-486-1
; Sequence 1, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-849-486-1

Query Match 51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQPKIWFNRRKPKWK 21
Db 1 RQIKIWFQNRMRKWK 16

RESULT 13
US-08-849-486-4
; Sequence 4, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-308-935-8

Query Match          51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQKIWFNNRRKPWK 21
      || |||| |||| ||||
DB      1 RQKIWFNNRRKPWK 16

RESULT 16
US-09-441-416A-6
; Sequence 6, Application US/09441416A
; Patent No. 6294518
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: SKOLNIK, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF
; TITLE OF INVENTION: CALPAIN
; FILE REFERENCE: 00398-140001
; CURRENT APPLICATION NUMBER: US/09/441,416A
; CURRENT FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: US 08/964,302
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-441-416A-6

Query Match          51.9%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQKIWFNNRRKPWK 21
      || |||| |||| ||||
DB      1 RQKIWFNNRRKPWK 16

RESULT 17
US-09-296-089-33
; Sequence 33, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
; FILE REFERENCE: 100086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-296-089-33

Query Match          51.9%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQKIWFNNRRKPWK 21

```


US-09-057-363C-47
; Sequence 47, Application US/09057363C
; Patent No. 6551994
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING THE
; INTERACTION BETWEEN ALPHA-CATENIN AND BETA-CATENIN
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,363C
; FILING DATE: 08-Apr-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 692-6031
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-057-363C-47

Query Match 51.9%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RPKIWFNNRRKPKWK 21
DB 1 RQIKWIFQNRRMKWK 16

RESULT 23
US-09-043-560B-3
; Sequence 3, Application US/09043560B
; Patent No. 6569833
; GENERAL INFORMATION:
; APPLICANT: Fahraeus, Robin
; APPLICANT: Lane, David P.
; TITLE OF INVENTION: Cyclin Dependent Kinase Binding Compounds
; FILE REFERENCE: CCI-0030S
; CURRENT APPLICATION NUMBER: US/09/043,560B
; CURRENT FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-043-560B-3

Query Match 51.9%; Score 67; DB 4; Length 16;

Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 6 RPKIWFNNRRKPKWK 21
DB 1 RQIKWIFQNRRMKWK 16
RESULT 24
US-09-648-400A-29
; Sequence 29, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; FILE REFERENCE: 019801-000210US
; CURRENT APPLICATION NUMBER: US/09/648,400A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Antennapedia
; OTHER INFORMATION: homeodomain, Antennapedia-43-58
US-09-648-400A-29

Query Match 51.9%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RPKIWFNNRRKPKWK 21
DB 1 RQIKWIFQNRRMKWK 16

RESULT 25
US-09-227-652B-4
; Sequence 4, Application US/09227652B
; Patent No. 6610495
; GENERAL INFORMATION:
; APPLICANT: TVM Telethon Institute for Child Health Research
; TITLE OF INVENTION: PEPTIDE DETECTION METHOD
; FILE REFERENCE: 1991209/MRO-ECT
; CURRENT APPLICATION NUMBER: US/09/227,652B
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 60/070989
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Penetratin 16-mer.
US-09-227-652B-4

Query Match 51.9%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RPKIWFNNRRKPKWK 21

Db 1 RQIKIWEQNRRMOKWK 16

Search completed: May 24, 2004, 17:29:50
Job time : 15.0405 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:19:01 ; Search time 39.4459 Seconds

(without alignments)
148.491 Million cell updates/sec

Title: US-09-977-349-4

Perfect score: 129

Sequence: 1 CSSCIQPKIWFNRRXPKK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	76.7	16	12	US-10-432-291-7
2	99	76.7	16	12	US-10-136-187-22
3	99	76.7	16	14	US-10-136-738-14
4	99	76.7	27	12	US-10-432-291-8
5	93	72.1	16	12	US-10-261-161-28
6	85	65.9	16	9	US-09-981-286A-5
7	82.5	64.0	17	9	US-09-854-204-59
8	75	58.1	16	9	US-09-854-204-58
9	75	58.1	16	9	US-09-981-286A-4
10	75	58.1	16	12	US-10-261-161-27
11	75	58.1	16	14	US-10-136-738-13
12	73	56.6	29	15	US-10-176-419A-4
13	69	53.5	17	9	US-09-854-204-22
14	68	52.7	115	9	US-09-925-299-1169
15	68	52.7	115	10	US-09-925-299-1169

67	51.9	16	8	US-08-610-220A-9	Sequence 9, Appli
67	51.9	16	9	US-09-214-371-43	Sequence 43, Appli
67	51.9	16	9	US-09-780-070-38	Sequence 38, Appli
67	51.9	16	9	US-09-150-623-9	Sequence 9, Appli
67	51.9	16	9	US-09-731-023A-10	Sequence 10, Appli
67	51.9	16	9	US-09-854-204-1	Sequence 1, Appli
67	51.9	16	9	US-09-854-204-56	Sequence 56, Appli
67	51.9	16	9	US-09-900-147-8	Sequence 8, Appli
67	51.9	16	9	US-09-792-480-29	Sequence 29, Appli
67	51.9	16	9	US-09-785-802A-2	Sequence 2, Appli
67	51.9	16	9	US-09-785-802A-5	Sequence 5, Appli
67	51.9	16	9	US-09-902-432-32	Sequence 32, Appli
67	51.9	16	9	US-09-953-031A-10	Sequence 10, Appli
67	51.9	16	9	US-09-981-286A-3	Sequence 3, Appli
67	51.9	16	10	US-09-962-967A-6	Sequence 6, Appli
67	51.9	16	10	US-09-912-414-6	Sequence 6, Appli
67	51.9	16	10	US-09-775-052-54	Sequence 54, Appli
67	51.9	16	10	US-09-295-189-4	Sequence 4, Appli
67	51.9	16	10	US-09-965-876A-1	Sequence 1, Appli
67	51.9	16	11	US-09-933-780C-2	Sequence 2, Appli
67	51.9	16	12	US-10-286-964-3	Sequence 3, Appli
67	51.9	16	12	US-10-432-291-6	Sequence 6, Appli
67	51.9	16	12	US-10-357-826A-2	Sequence 2, Appli
67	51.9	16	12	US-10-603-409-12	Sequence 12, Appli
67	51.9	16	12	US-09-748-063-3	Sequence 3, Appli
67	51.9	16	12	US-09-779-791A-3	Sequence 3, Appli
67	51.9	16	12	US-10-083-960-29	Sequence 29, Appli
67	51.9	16	12	US-10-144-549-1	Sequence 1, Appli
67	51.9	16	12	US-10-136-187-21	Sequence 21, Appli
67	51.9	16	12	US-10-261-161-1	Sequence 1, Appli
67	51.9	16	12	US-10-261-161-26	Sequence 26, Appli
67	51.9	16	13	US-10-024-935-12	Sequence 12, Appli
67	51.9	16	13	US-10-007-363-3	Sequence 3, Appli
67	51.9	16	14	US-10-071-512A-2	Sequence 2, Appli
67	51.9	16	14	US-10-239-804-3	Sequence 3, Appli
67	51.9	16	14	US-10-077-555-3	Sequence 3, Appli
67	51.9	16	14	US-10-209-421-29	Sequence 29, Appli
67	51.9	16	14	US-10-229-915-2	Sequence 2, Appli
67	51.9	16	14	US-10-185-084-3	Sequence 3, Appli
67	51.9	16	14	US-10-252-012-5	Sequence 5, Appli
67	51.9	16	14	US-10-075-869-19	Sequence 19, Appli
67	51.9	16	14	US-10-013-815-19	Sequence 19, Appli
67	51.9	16	14	US-10-136-738-10	Sequence 10, Appli
67	51.9	16	14	US-10-210-660-1	Sequence 1, Appli
67	51.9	16	14	US-10-210-660-25	Sequence 25, Appli
67	51.9	16	14	US-10-156-570A-21	Sequence 21, Appli
67	51.9	16	14	US-10-201-394A-14	Sequence 14, Appli
67	51.9	16	14	US-10-017-672-11	Sequence 11, Appli
67	51.9	16	14	US-10-201-389A-14	Sequence 14, Appli
67	51.9	16	14	US-10-161-051-1	Sequence 1, Appli
67	51.9	16	14	US-10-358-365-10	Sequence 10, Appli
67	51.9	16	14	US-10-061-607A-2	Sequence 2, Appli
67	51.9	16	14	US-10-405-339-44	Sequence 44, Appli
67	51.9	16	15	US-10-366-493-19	Sequence 19, Appli
67	51.9	16	15	US-10-444-662-2	Sequence 2, Appli
67	51.9	16	15	US-10-185-593-2	Sequence 2, Appli
67	51.9	16	15	US-10-413-160-38	Sequence 38, Appli
67	51.9	16	15	US-10-462-138-10	Sequence 10, Appli
67	51.9	16	15	US-10-369-226-47	Sequence 47, Appli
67	51.9	16	15	US-10-353-678-2	Sequence 2, Appli
67	51.9	16	15	US-10-357-529-8	Sequence 8, Appli
67	51.9	17	9	US-09-854-204-19	Sequence 19, Appli
67	51.9	17	9	US-09-854-204-27	Sequence 27, Appli
67	51.9	17	9	US-09-854-204-31	Sequence 31, Appli
67	51.9	17	9	US-09-854-204-32	Sequence 32, Appli
67	51.9	17	9	US-09-785-802A-3	Sequence 3, Appli
67	51.9	17	11	US-09-933-780C-21	Sequence 21, Appli
67	51.9	17	12	US-10-602-303-3	Sequence 3, Appli
67	51.9	17	12	US-10-421-548-8	Sequence 8, Appli
67	51.9	17	13	US-10-007-761-8	Sequence 8, Appli
67	51.9	17	14	US-10-209-421-30	Sequence 30, Appli
67	51.9	17	14	US-10-229-915-1	Sequence 1, Appli
67	51.9	17	14	US-10-210-660-17	Sequence 17, Appli

89 67 51.9 17 14 US-10-210-660-20 Sequence 20, Appl
 90 67 51.9 17 14 US-10-210-660-22 Sequence 22, Appl
 91 67 51.9 17 14 US-10-210-660-27 Sequence 27, Appl
 92 67 51.9 17 15 US-10-372-003A-29 Sequence 29, Appl
 93 67 51.9 17 15 US-10-428-280-15 Sequence 15, Appl
 94 67 51.9 17 15 US-10-421-503-66 Sequence 66, Appl
 95 67 51.9 18 9 US-09-785-802A-14 Sequence 14, Appl
 96 67 51.9 19 9 US-09-949-474-7 Sequence 7, Appl
 97 67 51.9 19 14 US-10-118-079-45 Sequence 45, Appl
 98 67 51.9 19 14 US-10-210-660-23 Sequence 23, Appl
 99 67 51.9 19 15 US-10-407-449-20 Sequence 20, Appl
 100 67 51.9 20 9 US-09-854-204-63 Sequence 63, Appl

ALIGNMENTS

RESULT 1
 US-10-432-291-7
 ; Sequence 7, Application US/10432291
 ; Publication No. US20040029281A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Centre National de la Recherche Scientifique (CNRS)
 ; APPLICANT: Joliot, Alain
 ; APPLICANT: Dupont, Edmond
 ; APPLICANT: Prochiantz, Alain
 ; TITLE OF INVENTION: Carrier vectors through an epithelium with tight junctions
 ; FILE REFERENCE: 45636-5067-US
 ; CURRENT APPLICATION NUMBER: US/10/432,291
 ; CURRENT FILING DATE: 2003-05-20
 ; PRIOR APPLICATION NUMBER: PCT/FR01/03631
 ; PRIOR FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: FR 00/14945
 ; PRIOR FILING DATE: 2000-11-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: penetratin sequence for transport vectors
 US-10-432-291-7

Query Match 76.7%; Score 99; DB 12; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
 |||||
 Db 1 RQPKWFFNRRKPKWK 16

RESULT 2
 US-10-136-187-22
 ; Sequence 22, Application US/10136187
 ; Publication No. US20030203865A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harvie, Pierrot
 ; APPLICANT: Paul, Ralph
 ; APPLICANT: Cudmore, Sally
 ; APPLICANT: O'Mahony, Daniel J.
 ; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
 ; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
 ; FILE REFERENCE: 226272005300
 ; CURRENT APPLICATION NUMBER: US/10/136,187
 ; CURRENT FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: US 60/287,786
 ; PRIOR FILING DATE: 2001-04-30
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 22
 ; LENGTH: 16

; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Construct
 US-10-136-187-22

Query Match 76.7%; Score 99; DB 12; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
 |||||
 Db 1 RQPKWFFNRRKPKWK 16

RESULT 3
 US-10-136-738-14
 ; Sequence 14, Application US/10136738
 ; Publication No. US2003010886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Finn, John
 ; APPLICANT: MacLachlan, Ian
 ; APPLICANT: Protiva Biopharmaceutics Inc.
 ; TITLE OF INVENTION: Autogene Nucleic Acids Encoding a
 ; TITLE OF INVENTION: Secretatable RNA Polymerase
 ; FILE REFERENCE: 020801-000310US
 ; CURRENT APPLICATION NUMBER: US/10/136,738
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/287,974
 ; PRIOR FILING DATE: 2001-04-30
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 14
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Antennapedia homeodomain third helix (residues
 ; OTHER INFORMATION: 43-58), 3-Pro secretion domain
 US-10-136-738-14

Query Match 76.7%; Score 99; DB 14; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
 |||||
 Db 1 RQPKWFFNRRKPKWK 16

RESULT 4
 US-10-432-291-8
 ; Sequence 8, Application US/10432291
 ; Publication No. US20040029281A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Centre National de la Recherche Scientifique (CNRS)
 ; APPLICANT: Joliot, Alain
 ; APPLICANT: Dupont, Edmond
 ; APPLICANT: Prochiantz, Alain
 ; TITLE OF INVENTION: Carrier vectors through an epithelium with tight junctions
 ; FILE REFERENCE: 45636-5067-US
 ; CURRENT APPLICATION NUMBER: US/10/432,291
 ; CURRENT FILING DATE: 2003-05-20
 ; PRIOR APPLICATION NUMBER: PCT/FR01/03631
 ; PRIOR FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: FR 00/14945
 ; PRIOR FILING DATE: 2000-11-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 8
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Artificial

FEATURE:
; OTHER INFORMATION: nuclear export and penetratin sequence for transport vectors
US-10-432-291-8

Query Match 76.7%; Score 99; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
Db 12 RQPKWFFNRRKPKWK 27

RESULT 5
US-10-261-161-28
; Sequence 28, Application US/102611161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; TITLE OF INVENTION: Transfer (FRET) Assays for Clostridial Toxins
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-261-161-28

Query Match 72.1%; Score 93; DB 12; Length 16;
Best Local Similarity 93.8%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
Db 1 RQPKWFFNRRKPKWK 16

RESULT 6
US-09-981-286A-5
; Sequence 5, Application US/09981286A
; Publication No. US20020192799A1
; GENERAL INFORMATION:
; APPLICANT: Watowich, Stanley J.
; APPLICANT: Weaver, Scott C.
; APPLICANT: Davey, Robert A.
; TITLE OF INVENTION: Drug Discovery Methods
; FILE REFERENCE: 265.00260101
; CURRENT APPLICATION NUMBER: US/09/981,286A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/240,187
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-5

Query Match 65.9%; Score 85; DB 9; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 RQPKWFFNRRKPKWK 21
Db 1 RQPKWFFNRRKPKWK 16

RESULT 7
US-09-854-204-59
; Sequence 59, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-59

Query Match 64.0%; Score 82.5; DB 9; Length 17;
Best Local Similarity 88.2%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 6 RQPKWFFNRRKPKWK 21
Db 1 RQPKWFFNRRKPKWK 17

RESULT 8
US-09-854-204-58
; Sequence 58, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-58

Query Match 58.1%; Score 75; DB 9; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPKWK 21
DB 1 RQKIWFPPNRRKPKWK 16

RESULT 9
US-09-981-286A-4
; Sequence 4, Application US/09981286A
; Publication No. US20020192799A1
; GENERAL INFORMATION:
; APPLICANT: Watowich, Stanley J.
; APPLICANT: Weaver, Scott C.
; APPLICANT: Davey, Robert A.
; TITLE OF INVENTION: Drug Discovery Methods
; FILE REFERENCE: 265.00260101
; CURRENT APPLICATION NUMBER: US/09/981,286A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/240,187
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-4

Query Match 58.1%; Score 75; DB 9; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPKWK 21
DB 1 RQKIWFPPNRRKPKWK 16

RESULT 10
US-10-261-161-27
; Sequence 27, Application US/10261161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 16
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-261-161-27

Query Match 58.1%; Score 75; DB 12; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPKWK 21
DB 1 RQKIWFPPNRRKPKWK 16

RESULT 11
US-10-136-738-13
; Sequence 13, Application US/10136738
; Publication No. US20030108896A1
; GENERAL INFORMATION:
; APPLICANT: Finn, John
; APPLICANT: MacLachlan, Ian
; APPLICANT: Protiva Biotherapeutics Inc.
; TITLE OF INVENTION: Autogene Nucleic Acids Encoding a
; TITLE OF INVENTION: Secretatable RNA Polymerase
; FILE REFERENCE: 020801-000310US
; CURRENT APPLICATION NUMBER: US/10/136,738
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,974
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antennapedia homeodomain third helix (residues
; OTHER INFORMATION: 43-58), Pro50 secretion domain
US-10-136-738-13

Query Match 58.1%; Score 75; DB 14; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPKWK 21
DB 1 RQKIWFPPNRRKPKWK 16

RESULT 12
US-10-176-419A-4
; Sequence 4, Application US/10176419A
; Publication No. US20040006203A1
; GENERAL INFORMATION:
; APPLICANT: Maier, Martin A.
; APPLICANT: Guzaev, Andrei P.
; APPLICANT: Manoharan, Muthiah
; TITLE OF INVENTION: Method For Solid Phase Synthesis Of PNA Conjugates Using Branched
; TITLE OF INVENTION: Bridging Units Involving Orthogonal Protecting Groups
; FILE REFERENCE: ISIS057
; CURRENT APPLICATION NUMBER: US/10/176,419A
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide sequence
US-10-176-419A-4

Query Match 56.6%; Score 73; DB 15; Length 29;

```
Best Local Similarity 51.9%; Pred. No. 0.01;
Matches 14; Conservative 2; Mismatches 5; Indels 6; Gaps 1;

Qy 1 CSSCI-----RQKIWFNNRRKPKWK 21
   |||||
Db 3 CTTCCCKKRRQKIWFNNRRKPKWK 29

RESULT 13
US-09-854-204-22
; Sequence 22, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)-
; OTHER INFORMATION: bala
; NAME/KEY: MOD_RES
; LOCATION: (17)
; OTHER INFORMATION: AMIDATION
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-22

Query Match 53.5%; Score 69; DB 9; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKIWFNNRRKPKWK 21
   |||||
Db 2 RQKIWFNNRRKPKWK 17

RESULT 14
US-09-925-299-1169
; Sequence 1169, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1169
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1169

Query Match 52.7%; Score 68; DB 9; Length 115;
Best Local Similarity 75.0%; Pred. No. 0.16;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKIWFNNRRKPKWK 21
   |||||
Db 64 RQKIWFNNRRKPKWK 79

RESULT 15
US-09-925-299-1169
; Sequence 1169, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1169
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1169

Query Match 52.7%; Score 68; DB 10; Length 115;
Best Local Similarity 75.0%; Pred. No. 0.16;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RQKIWFNNRRKPKWK 21
   |||||
Db 64 RQKIWFNNRRKPKWK 79

RESULT 16
US-08-610-220A-9
; Sequence 9, Application US/08610220A
; Publication No. US20030099638A1
; GENERAL INFORMATION:
; APPLICANT: Troy, Carol M.
; TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/610,220A
; FILING DATE: MAR-04-1996
; CLASSIFICATION: 424
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 48332/JFW/JWL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-610-220A-9

Query Match 51.9%; Score 67; DB 8; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRKPKWK 21
DB 1 RQKIWFQNRMRKWK 16

RESULT 17
US-09-214-371-43
; Sequence 43, Application US/09214371B
; Patent No. US20010018511AI
; GENERAL INFORMATION:
; APPLICANT: Lane, David
; APPLICANT: Bottger, Volker
; APPLICANT: Bottger, Angelica
; APPLICANT: Pickaley, Stephen
; APPLICANT: Chene, Patrick
; APPLICANT: Hochkeppel, Heinz-Kurt
; APPLICANT: Garcia-Echeverria, Carlos
; APPLICANT: Furet, Pascal
; TITLE OF INVENTION: Inhibitors of the Interaction of p53 and MDM2
; FILE REFERENCE: 4-20937/A/PCT
; CURRENT APPLICATION NUMBER: US/09/214,371B
; CURRENT FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: PCT/EP97/03549
; PRIOR FILING DATE: 1997-07-04
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
; US-09-214-371-43

Query Match 51.9%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRKPKWK 21
DB 1 RQKIWFQNRMRKWK 16

RESULT 18
US-09-780-070-38
; Sequence 38, Application US/09780070
; Patent No. US20020009752A1
; GENERAL INFORMATION:
; APPLICANT: Burke, James
; APPLICANT: Strittmatter, Warren
; APPLICANT: Nagai, Yoshitaka
; TITLE OF INVENTION: COMPOUNDS THAT SELECTIVELY BIND TO EXPANDED POLYGLUTAMINE REPEAT
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR

```

```
RESULT 20
US-09-731-023A-10
; Sequence 10, Application US/09731023A
; Patent No. US20020077283A1
; GENERAL INFORMATION:
; APPLICANT: Sessa, William
; TITLE OF INVENTION: Caveolin Peptides and Their Use as Therapeutics
; FILE REFERENCE: 44574-5076-US
; CURRENT APPLICATION NUMBER: US/09/731,023A
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/231,327
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; NAME/KEY: DOMAIN
; LOCATION: (1)..(16)
; OTHER INFORMATION: Homeodomain, internalization sequence
US-09-731-023A-10

Query Match          51.9%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQPKIWFNNRRKPKWK 21
      ||||| ||||| |||
Db      1 RQIKIWFQNRKPKWK 16

RESULT 21
US-09-854-204-1
; Sequence 1, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (12)
; OTHER INFORMATION: Xaa is norleucine ornithine
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-56

Query Match          51.9%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQPKIWFNNRRKPKWK 21
      ||||| ||||| |||
Db      1 RQIKIWFQNRKPKWK 16

RESULT 23
US-09-900-147-8
; Sequence 8, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laeantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-05-27
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-05-27
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

Query Match          51.9%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 RQPKIWFNNRRKPKWK 21
      ||||| ||||| |||
Db      1 RQIKIWFQNRKPKWK 16
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:24:46 ; Search time 54 Seconds

(without alignments)

41.859 Million cell updates/sec

Title: US-09-977-349-5

Perfect score: 43

Sequence: 1 PSYVNVQN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

A Geneseqp_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	12	5	Abg70971 Human SHC
2	43	100.0	15	5	Aae18721 Major his
3	43	100.0	473	2	Aar84637 Shc prote
4	43	100.0	473	2	Aaw14005 Human SHC
5	43	100.0	473	2	Aay13448 Amino aci
6	43	100.0	474	2	Aar97243 SHC phosph
7	43	100.0	583	3	Aab19072 Amino aci
8	43	100.0	583	4	Aab90782 Human she
9	42	97.7	469	2	Aay13449 Amino aci
10	38	88.4	12	5	Aau10435 Human SHC
11	37	86.0	9	2	Aar69827 Peptide c
12	37	86.0	9	2	Aaw46895 SHC phosph
13	37	86.0	9	5	Aau10467 Nucleic a
14	37	86.0	9	5	Aau10458 Nucleic a
15	37	86.0	9	5	Aae15236 SHC phosph
16	36	83.7	61	5	Abp33784 Human ORF
17	35	81.4	8	2	Aaw78664 SH2 domai
18	35	81.4	9	2	Aaw78663 SH2 domai
19	35	81.4	16	2	Aaw78661 SH2 domai
20	35	81.4	282	6	Ada35471 Acinetoba
21	34	79.1	187	6	Abz52862 Protein s
22	34	79.1	251	7	Adc94754 E. faeciu
23	34	79.1	273	6	Ada34556 Acinetoba
24	34	79.1	471	6	Abu24418 Protein e
25	33	76.7	144	4	Aam95180 Human rep

26	33	76.7	144	4	ABB95880
27	33	76.7	215	4	AAB92664
28	33	76.7	215	4	AAB88321 Human mem
29	33	76.7	256	7	ADB65101 Human pro
30	33	76.7	267	3	AAy94890 Human pro
31	33	76.7	267	4	AAB68874 Human REC
32	33	76.7	267	4	AAB65010 Human sec
33	33	76.7	322	5	ABP69845 Human pol
34	33	76.7	577	6	ABM73375 Staphyloc
35	33	76.7	732	4	ABb71917 Drosophil
36	33	76.7	880	3	AAy53621 Amino aci
37	33	76.7	4345	4	ABB66417 Drosophil
38	32	74.4	73	5	ABJ03721 Human ova
39	32	74.4	82	5	ABY79310 Human ova
40	32	74.4	82	5	ABJ03776 Human ova
41	32	74.4	141	3	AAy73415 Human sec
42	32	74.4	200	4	AAU23734 Novel hum
43	32	74.4	200	4	AAU23196 Novel hum
44	32	74.4	503	6	ABU36317 Protein e
45	32	74.4	538	4	AAm40339 Human pol
46	32	74.4	578	4	AAm40338 Human pol
47	32	74.4	578	5	ABa47734 Human Ren
48	32	74.4	578	6	AAc19623 Human nox
49	32	74.4	578	7	ADD48774 Human pro
50	32	74.4	622	5	ABP74053 Candida a
51	32	74.4	1328	6	AAE37961 Human kin
52	32	74.4	1360	4	AAU11802 Human pro
53	32	74.4	1360	5	AAE25096 Human kin
54	31	72.1	39	4	AAm61411 Human bra
55	31	72.1	51	4	ABb42690 Peptide #
56	31	72.1	51	4	AAm36503 Peptide #
57	31	72.1	51	4	AAm76395 Human bon
58	31	72.1	51	4	AAm63580 Human bra
59	31	72.1	66	4	AAu86498 Novel hum
60	31	72.1	66	7	ADb59832 Connectiv
61	31	72.1	144	2	AAy20369 Human mic
62	31	72.1	160	2	AAW18655 Human mic
63	31	72.1	239	3	AAy74795 Neisseria
64	31	72.1	239	3	AAy74794 Neisseria
65	31	72.1	273	6	ABu22950 Protein e
66	31	72.1	273	6	ABu45279 Protein e
67	31	72.1	273	6	ABU17005 Protein e
68	31	72.1	273	6	ABU38161 Protein e
69	31	72.1	273	6	ABU37457 Protein e
70	31	72.1	274	6	ABU50030 Protein e
71	31	72.1	274	6	ABU28499 Protein e
72	31	72.1	274	6	ABU48218 Protein e
73	31	72.1	274	6	ABU27888 Protein e
74	31	72.1	274	6	ABU39133 Protein e
75	31	72.1	274	6	ABU40640 Protein e
76	31	72.1	275	6	ABM67552 Phototrab
77	31	72.1	275	6	ABU32089 Protein e
78	31	72.1	276	6	ABU33296 Protein e
79	31	72.1	280	4	ABG26067 Novel hum
80	31	72.1	285	6	ADA35070 Acinetoba
81	31	72.1	304	6	ABU30642 Protein e
82	31	72.1	314	4	AAE02719 Zea may
83	31	72.1	337	6	ABM69077 Phototrab
84	31	72.1	347	2	AAr61990 Porphyrom
85	31	72.1	348	2	AAr61991 Porphyrom
86	31	72.1	348	2	AAr61989 Porphyrom
87	31	72.1	353	6	ABP78717 N. gonorr
88	31	72.1	388	3	AAg28060 Arabidops
89	31	72.1	392	3	AAg48209 Arabidops
90	31	72.1	405	3	AAg28059 Arabidops
91	31	72.1	408	7	ADc94286 E. faeciu
92	31	72.1	410	3	AAg48208 Arabidops
93	31	72.1	442	4	ABG14388 Novel hum
94	31	72.1	472	2	AAW37243 Globodera
95	31	72.1	484	7	ADE10039 Novel pro
96	31	72.1	491	4	ABG26065 Novel hum
97	31	72.1	803	4	ABG30222 Novel hum
98	31	72.1	894	2	AAr42963 Human MTP

```

99 31 72.1 894 2 AAR50007 Human WTP
100 31 72.1 894 7 ADB68723 Mouse mic
101 31 72.1 894 7 ADB68722 Human mic
102 31 72.1 923 4 ABG20904 Novel hum
103 31 72.1 1009 5 AAO20949 Protein o
104 31 72.1 1432 2 AAY14517 Human WRN
105 31 72.1 1432 5 AAM50935 Human WR
106 31 72.1 1432 5 ADC84515 GenPept a
107 30 69.8 26 ABG80398 Plant FD/
108 30 69.8 61 ABP39729 Staphyloc
109 30 69.8 154 ABB61210 Drosophil
110 30 69.8 225 3 AAY95940 Porcine a
111 30 69.8 263 5 ABG93168 S. cerevi
112 30 69.8 271 5 AAM51458 Human mic
113 30 69.8 284 5 ABP30424 Streptoco
114 30 69.8 284 6 ABU01353 S. pneumo
115 30 69.8 285 3 AAY81540 Streptoco
116 30 69.8 285 5 ABP25457 Streptoco
117 30 69.8 290 4 ABB61035 Drosophil
118 30 69.8 300 2 AAY41759 Human PRO
119 30 69.8 300 3 AAB44315 Human PRO
120 30 69.8 300 4 AAU29089 Human PRO
121 30 69.8 300 4 AAB19930 Human PRO
122 30 69.8 300 6 ABUS8465 Human PRO
123 30 69.8 300 6 ABUS8013 Novel hum
124 30 69.8 300 6 ABUS84328 Human sec
125 30 69.8 300 6 ABR66202 Human sec
126 30 69.8 300 6 ABR65592 Human sec
127 30 69.8 300 6 ABUS9532 Human PRO
128 30 69.8 300 6 ABUS2771 Human PRO
129 30 69.8 300 6 ABUS9892 Novel hum
130 30 69.8 300 6 ABR68141 Human sec
131 30 69.8 300 6 ABUS96194 Novel hum
132 30 69.8 300 6 ABUS2625 Human sec
133 30 69.8 300 6 ABO08702 Human sec
134 30 69.8 300 6 ABO02754 Human sec
135 30 69.8 300 6 ABR74908 Human sec
136 30 69.8 300 6 ABR94670 Human sec
137 30 69.8 300 6 ABO25261 Novel hum
138 30 69.8 300 6 ABUS5643 Human PRO
139 30 69.8 300 6 ABUS9803 Novel hum
140 30 69.8 300 6 ABUS98018 Novel hum
141 30 69.8 300 6 ABUS91724 Novel hum
142 30 69.8 300 6 ABUS72267 Novel hum
143 30 69.8 300 6 ABUS9417 Human PRO
144 30 69.8 300 6 ABUS6258 Human sec
145 30 69.8 300 6 ABUS7471 Human sec
146 30 69.8 300 6 ABUS80499 Human PRO
147 30 69.8 300 6 ABR99417 Human sec
148 30 69.8 300 6 ABR98807 Human sec
149 30 69.8 300 6 ABO16330 Human sec
150 30 69.8 300 6 ABR92230 Human sec

```

ALIGNMENTS

```

RESULT 1
ABG70971
ID ABR70971 standard; peptide; 12 AA.
XX
AC ABR70971;
XX
DT 04-DEC-2002 (first entry)
XX
DE Human SHC peptide.
XX
KW Human; radiotherapeutic agent precursor; targeting peptide; tumour;
KW phosphorylation peptide; linker; Src homology region 2 domain; SH_2;
KW infectious lesion; breast; colon; rectum; prostate; ovary; testes;
KW pathological lesion; chest; abdomen; pelvis; skin; cardiovascular lesion;
KW myocardial infarct; atherosclerotic plaque; clot; thrombosis;
KW pulmonary embolism; inflammatory lesion; hyperplasia; SHC.

```

```

XX OS Homo sapiens.
XX US6440386-B1.
XX 27-AUG-2002.
XX PD
XX AAM50935 Human WR
XX PF 06-JUL-1999; 99US-00347926.
XX PR 06-JUL-1998; 98US-0091736P.
XX (IMMU-) IMMUNOMEDICS INC.
XX Leung S;
XX WPI; 2002-722167/78.
XX Radioactive agent precursor, useful in manufacture of in vivo
XX radiotherapeutic agents, comprises protein containing targeting peptide,
XX phosphorylation peptide, linker and Src homology region 2 domain.
XX Disclosure; Col 8; 13pp; English.
XX The present invention relates to a new radiotherapeutic agent precursor
XX comprising a protein containing a targeting peptide, a phosphorylation
XX peptide, a linker and Src homology region 2 domain (SH_2). The invention
XX is useful for preparing radiotherapeutic agents, useful for the treatment
XX of tumour and an infectious lesion. The lesions include infectious
XX lesions of breast, colon, rectum, prostate, ovary and testes;
XX pathological lesions affecting chest, abdomen and pelvis and skin; and
XX cardiovascular lesions such as myocardial infarct, atherosclerotic
XX plaque, clot, thrombosis, pulmonary embolism, inflammatory lesions and
XX hyperplasia. The linker in the radiotherapeutic agent is long and fold
XX sufficiently flexible to allow the SH_2 domain to make a turn and fold
XX back on the phosphorylated substrate. Thus, the radiotherapeutic agent
XX exhibits improved stability in vivo, by avoiding the ready
XX phosphorylation, while maintaining the binding ability of the agent. The
XX present amino acid sequence represents a human peptide that contains an
XX SH_2 domain as described in the invention
XX
XX Sequence 12 AA;
XX
XX Query Match 100.0%; Score 43; DB 5; Length 12;
XX Best Local Similarity 100.0%; Pred. NO. 0.028;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PSYVNVQN 8
XX DB 5 PSYVNVQN 12
XX
XX RESULT 2
XX AAE18721
XX ID AAE18721 standard; peptide; 15 AA.
XX AC AAE18721;
XX DT 17-MAY-2002 (first entry)
XX DE
XX Major histocompatibility complex class II immunogenic peptide #27.
XX Major histocompatibility complex; MHC; cytostatic; vaccine; immunogen;
XX cytotoxic T lymphocyte; CTL; carcinoma; tumour; cytokine; immunotherapy;
XX cancer; prostate; colorectal; ovarian; lung.
XX Unidentified.
XX WO200197827-A1.
XX 27-DEC-2001.
XX 15-JUN-2001; 2001WO-US019207.
XX

```

PR 16-JUN-2000; 2000US-0212165P.
PR 16-JUN-2000; 2000US-0212213P.

XX (ARGO-) ARGONEX PHARM.
XX (UYVI-) UNIV VIRGINIA PATENT FOUND.

XX White F, Hunt DF, Shabanowitz J, Ross M;
XX WPI; 2002-179540/23.

XX Immunogen, such as a major histocompatibility complex polypeptide, over-
PT expressed on prostate cancer cells, useful for prevention, treatment and
PT diagnosis of cancer, preferably prostate cancer.

XX Claim 1; Page 23; 65pp; English.

XX The invention relates to an immunogen such as a major histocompatibility
CC complex (MHC) polypeptide over-expressed on prostate cancer cells. The
CC immunogen comprises an oligopeptide, a peptide mimetic of the oligo
CC nucleotide or an oligopeptide having a sequence differing by up to three
CC conservative amino acid substitutions. The immunogen is useful for
CC inducing a cytotoxic T lymphocyte (CTL) in vitro that is specific for a
CC prostate carcinoma. The induced CTL is useful for treating a subject with
CC cancer, preferably prostate cancer, or for treating a cancer-afflicted
CC subject characterized by tumor cells expressing any class I MHC
CC molecule and a gene coding for an epitopic sequence, by administering the
CC induced CTLs in an amount sufficient to destroy the tumor cells through
CC direct lysis or to effect the destruction of the tumor cells indirectly
CC through the elaboration of cytokines. The immunogen is useful for
CC prevention, treatment and diagnosis of cancer, preferably prostate
CC cancer, colorectal cancer, ovarian cancer or lung cancer, as diagnostic
CC markers of cancer, as screening agents, to stimulate the production of
CC antibodies for use in passive immunotherapy, and as reagents in processes
CC such as affinity chromatography. The immunogen is also for development of
CC vaccines and as therapeutic agents for stimulation of the immune system
CC for killing prostate cancer cells. The present sequence is MHC class II
CC immunogenic peptide derived from SHC transforming protein (p66shc)

XX Sequence 15 AA;

Query Match 100.0%; Score 43; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 2 PSYVNVQN 9

RESULT 3
AAR84637
ID AAR84637 standard; protein; 473 AA.

XX AAR84637;

DT 25-FEB-1996 (first entry)

XX Shc protein.

XX Shc; Grb2; BCR-ABL; tyrosine kinase; transformation; Ras; oncoprotein;
KW leukaemia.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Binding-site 317..320

FT /label= Grb2-binding_site

FT Modified-site 317

FT /label= Phosphorylation_site

XX CA2113494-A.

XX 15-JUL-1995.

XX 14-JAN-1994; 94CA-02113494.

XX 14-JAN-1994; 94CA-02113494.

XX (MOUN) MOUNT SINAI HOSPITAL CORP.
XX (TEXA) UNIV TEXAS.

XX Paul L, Pawson A, Arlinghaus R, Gish G, Liu J;

XX WPI; 1995-302931/40.
XX N-PSDB; AAT05112.

XX Detection of agents that modify BCR-ABL mediated transformation - useful
PT in treatment of leukaemia and other malignancies.

XX Disclosure; Page 50-51; 106pp; English.

XX The human Shc protein (AAR84637) has a Grb2 binding site useful for
CC screening compounds that affect BCL-ABL mediated transformation. Such
CC compounds have value in the treatment of chronic, acute myelogenous or
CC acute lymphocytic leukaemia

XX Sequence 473 AA;

Query Match 100.0%; Score 43; DB 2; Length 473;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8

Db 315 PSYVNVQN 322

RESULT 4

AAW14005

ID AAW14005 standard; protein; 473 AA.

XX AAW14005;

DT 24-JUN-1997 (first entry)

XX Human SHC protein.

XX SH2-containing inositol phosphatase; SHIP;
KW inositol polyphosphate 5-phosphatase; src homology domain 2; SH2 domain;
KW signal transduction; leukaemia; cancer; SHC.

XX Homo sapiens.

XX WO9712039-A2.

XX 03-APR-1997.

XX 27-SEP-1996; 96WO-CA000655.

XX 27-SEP-1995; 95US-0006063P.

XX 30-NOV-1995; 95US-000788P.

XX 09-APR-1996; 96US-0015217P.

XX 14-JUN-1996; 96US-00664962.

XX (KRYSTAL) KRYSTAL G.

XX Krystal G;

XX WPI; 1997-212898/19.

XX N-PSDB; AAT60305.

XX Inositol polyphosphate-5-phosphatase having SH2 domain - useful for
PT treating cancer and other conditions involving abnormal signalling.

XX Disclosure; Page 42-45; 89pp; English.

CC Human SHC (AAW14005) is a transforming protein with an Src homology
 CC domain 2 (SH2). It is implicated in mitogenic signal transduction. SHC is
 CC capable of binding novel murine and human SHP (SH2-containing inositol
 CC phosphatase) proteins (see also AAW14002-03). SHC, or its SH2 domain,
 CC phosphotyrosine binding recognition sequence, or region contg. Tyr-317,
 CC can be used in methods for assaying agonists and antagonists of SHP
 XX
 SQ Sequence 473 AA;

Query Match 100.0%; Score 43; DB 2; Length 473;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||
 Db 315 PSYVNVQN 322

RESULT 5
 AAY13448
 ID AAY13448 standard; protein; 473 AA.

XX AC AAY13448;

XX DT 26-JUL-1999 (first entry)

XX DE Amino acid sequence of human Shc.

XX KW Phosphotyrosine interaction domain; PI domain; modulator; APP;
 KW amyloid precursor protein; cytoplasmic domain; Alzheimer's disease;
 KW neurodegenerative disease; dementia; Shc.

XX OS Homo sapiens.

XX XX WO9221995-A1.

XX PD 06-MAY-1999.

XX PF 23-OCT-1998; 98WO-US022523.

XX PR 24-OCT-1997; 97US-00957660.

XX PR 17-APR-1998; 98US-00062085.

XX XX (UYRQ) UNIV ROCKEFELLER.

XX PI Sabo S, Buxbaum J, Greengard P;

XX XX WPI; 1999-326703/27.

XX DT Cell line for screening agents that inhibit processing of amyloid
 PT precursor protein.

XX PS Claim 4; Fig 3; 86pp; English.

XX CC The invention relates to modulators that affect the interaction of a
 CC polypeptide defining a PI (phosphotyrosine interaction) domain, or its
 CC conserved variants or fragments, with APP (amyloid precursor protein).
 CC The polypeptide defining a PI domain binds to the cytoplasmic domain of
 CC APP, regulating the secretion of APP fragments. By modulating this
 CC interaction, the modulator can inhibit or prevent APP processing and
 CC trafficking, and thus progress or onset of Alzheimer's disease. The
 CC modulators are used to treat and/or prevent neurodegenerative diseases in
 CC mammals, particularly Alzheimer's disease and dementia. Sequences
 CC AAY13448-461 represent examples of polypeptides defining a PI domain
 XX
 SQ Sequence 473 AA;

Query Match 100.0%; Score 43; DB 2; Length 473;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||

Db 315 PSYVNVQN 322
 RESULT 6
 AAR97243
 ID AAR97243 standard; protein; 474 AA.

XX AC AAR97243;

XX DT 21-AUG-1996 (first entry)

XX DE SHC phosphotyrosine binding domain.

XX KW Phosphotyrosine binding domain; SHC; cell signalling; atherosclerosis;
 KW inflammatory joint disease; psoriasis; restinosis; cancer; therapy;
 KW diagnosis.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers
 FT Domain 46..232
 FT /label= Phosphotyrosine_binding_domain

XX PN WO9617866-A1.

XX PD 13-JUN-1996.

XX PF 20-NOV-1995; 95WO-US015144.

XX PR 09-DEC-1994; 94US-00353550.

XX XX (REGC) UNIV CALIFORNIA.

XX PI Williams LT, Kavanaugh WM;

XX DR WPI; 1996-287115/29.

XX DR N-PSDB; AAT09252.

XX PT Isolated polypeptide(s) with phosphotyrosine binding domains - used to
 PT develop prods. for studying cell signalling pathways and for treating
 PT proliferative cell disorders.

XX PS Claim 18; Fig 5; 73pp; English.

XX CC An N-terminal portion (AAR97243) of the SHC protein contains a domain
 CC that is capable of binding to tyrosine-phosphorylated proteins in a
 CC manner similar to that of SH2 domains. This phosphotyrosine binding
 CC domain (PTB) was identified by deletion analysis of SHC. Recombinant PTB
 CC can be obtained by insertion of encoding DNA (see also AAR09252) into a
 CC vector and expression in host cells, pref. Escherichia coli or Sf9 insect
 CC cells. The PTB can be used to develop prods. for studying cell signalling
 CC pathways, or to screen potential PDB (ant)agonists, to determine whether
 CC a protein is a phosphorylated ligand of a PTB domain, or to block growth
 CC factor dependent stimulation of cells

XX SQ Sequence 474 AA;

Query Match 100.0%; Score 43; DB 2; Length 474;

Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0;

QY 1 PSYVNVQN 8
 |||||

Db 316 PSYVNVQN 323

RESULT 7
 AAB19072
 ID AAB19072 standard; protein; 583 AA.

XX AC AAB19072;

XX DT 08-FEB-2001 (first entry)

```

XX DE Amino acid sequence of a murine p66 protein.
XX DE
XX KW p66; stress resistance; aging; tumour suppression; p53; apoptosis;
XX KW cellular senescence; lung emphysema; myocardial infarction; stroke;
XX KW premature aging; Parkinson's disease; Alzheimer's disease; cancer;
XX KW diabetes.
XX OS Mus sp.
XX PN WO200056886-A1.
XX PD 28-SEP-2000.
XX PF 22-MAR-2000; 2000WO-GB001079.
XX PR 22-MAR-1999; 99GB-00006515.
XX PX (CANC-) CANCER RES VENTURES LTD.
XX PA Pellicci PG, Giorgio M, Migliaccio E, Lanfranccone L;
XX PI WPI; 2000-628262/60.
XX DR N-PSDB; AA96689.
XX CC The present sequence represents a p66 polypeptide. The specification
XX CC describes a p66 polypeptide, where at least 1 serine residue is absent or
XX CC replaced by a different amino acid residue. Targeted mutations of p66
XX CC gene induce stress resistance and prolong survival. The p66 gene is
XX CC involved in the process of aging and in tumour suppression. p66 is a
XX CC downstream target of p53 in the execution of the p53-mediated apoptosis
XX CC and cellular senescence. Manipulation of p66 can therefore be used to
XX CC restore p53 function in p53 null cells. Agonists of p66 may be used as
XX CC tumour suppressors. Any molecules that prevent p66 dephosphorylation or
XX CC kinases that induce p66 phosphorylation are potential tumour treatment
XX CC agents. Antibodies to p66 and antisense sequences to p66 are useful for
XX CC preparing a medicament useful for treating lung emphysema, myocardial
XX CC infarction, stroke, premature aging, cell senescence, Parkinson's
XX CC disease, Alzheimer's disease, cancers and diabetes
XX SQ Sequence 583 AA;
XX
XX Query Match 100.0%; Score 43; DB 3; Length 583;
XX Best Local Similarity 100.0%; Pred. No. 2.3;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PSYNNVQN 8
XX Db 425 PSYNNVQN 432
XX
XX RESULT 8
XX ID AAB90782 standard; protein; 583 AA.
XX AC AAB90782;
XX XX
XX DT 15-JUN-2001 (first entry)
XX DE Human shear stress-response protein SEQ ID NO: 64.
XX DE Human; shear stress-response protein; vascular disease; arteriosclerosis.
XX OS Homo sapiens.
XX PN WO200125427-A1.
XX PS
XX
XX Query Match 100.0%; Score 43; DB 3; Length 583;
XX Best Local Similarity 100.0%; Pred. No. 2.3;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PSYNNVQN 8
XX Db 425 PSYNNVQN 432
XX
XX RESULT 8
XX ID AAB90782 standard; protein; 583 AA.
XX AC AAB90782;
XX XX
XX DT 15-JUN-2001 (first entry)
XX DE Human shear stress-response protein SEQ ID NO: 64.
XX DE Human; shear stress-response protein; vascular disease; arteriosclerosis.
XX OS Homo sapiens.
XX PN WO200125427-A1.
XX PS
XX
XX Query Match 100.0%; Score 43; DB 4; Length 583;
XX Best Local Similarity 100.0%; Pred. No. 2.3;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PSYNNVQN 8
XX Db 425 PSYNNVQN 432
XX
XX RESULT 9
XX ID AAY13449 standard; protein; 469 AA.
XX AC AAY13449;
XX XX
XX DT 26-JUL-1999 (first entry)
XX DE Amino acid sequence of murine Shc.
XX DE
XX KW Phosphotyrosine interaction domain; PI domain; modulator; APP;
XX KW amyloid precursor protein; cytoplasmic domain; Alzheimer's disease;
XX KW neurodegenerative disease; dementia; Shc.
XX OS Mus sp.
XX PN WO9921995-A1.
XX PD 06-MAY-1999.
XX PF 23-OCT-1998; 98WO-US022523.
XX PR 24-OCT-1997; 97US-00957660.
XX PR 17-APR-1998; 98US-00062085.
XX PA (UVRQ ) UNIV ROCKEFELLER.
XX PI Sabo S, Buxbaum J, Greengard P;
XX DR WPI; 1999-326703/27.
XX CC Cell line for screening agents that inhibit processing of amyloid
XX CC precursor protein.
XX PS Claim 4; Fig 4; 86pp; English.

```

XX The invention relates to analysing the activity or level of one or more
CC proteins or enzymes, comprising providing a pool of substrates (e.g.
CC peptides, antibodies, binding domains or other molecules which act as
CC substrates or control substrates), each with a specific tag and
CC representing a substrate of one or more of the proteins or enzymes, or
CC substrates derived from these using the tagged substrates as substrates,
CC hybridising the pool of tagged substrates to an ordered array of specific,
CC and complementary tags immobilised on a surface, where the array
CC comprises different tags, at least some of which are control tags, each
CC tag is localised in a predetermined region of the surface and the density

XX New peptide(s) which bind to a cytosolic signal transducer - used for
 PT interfering with pathways leading to cell proliferation and motility, for
 PT treating neoplastic disease.
 XX Claim 10; Page 64; 19pp; English.
 XX This peptide is a tyrosine-contg. mol. representing a site of tyrosine
 CC phosphorylation. The peptides of the invention comprise a portion of the
 CC intracellular region of the human hepatocyte growth factor receptor
 CC (hGFR), and can generally reproduce potential recognition motifs for the
 CC SH2 domains of intracellular (cytosolic) signal inducers. The peptides
 CC can interfere with pathways leading to cell proliferation, movement and
 CC extracellular matrix invasion. They can be used to inhibit growth of
 CC neoplastic cells and to prevent metastatic spreading. (Updated on 25-MAR-
 CC 2003 to correct PN field.)
 XX Sequence 9 AA;
 SQ

Query Match 86.0%; Score 37; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSYVNVQ 7
 DB 3 PSYVNVQ 9

RESULT 12
 AAW46895
 ID AAW46895 standard; peptide; 9 AA.
 AC AAW46895;
 XX
 XX 19-JUN-1998 (first entry)
 DE SHC phosphopeptide capable of binding to the SH2 domain of Grb2.
 XX
 KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer.
 XX Unidentified.
 OS

Key Location/Qualifiers
 FT Modified-site 5
 FT /note= "phosphorylated"
 XX
 XX WO9802176-A1.
 XX 22-JAN-1998.
 XX
 XX 16-JUL-1997; 97WO-US012501.
 XX
 XX 16-JUL-1996; 96US-0021858P.
 XX (GEOU) UNIV GEORGETOWN.
 XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX King CR, Sastry L, Krag D, Oligino L;
 XX WPI; 1998-110340/10.
 DR
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX Disclosure; Page 9; 39pp; English.

XX The present sequence represents a SHC phosphopeptide that is capable of
 CC binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a signal
 CC transduction protein. The specification describes a non-phosphorylated
 CC peptide (AAW46896) that binds to the SH2 domain of Grb2 with affinity

CC similar to, or greater than, that of the present sequence. This non-
 CC phosphorylated peptide contains a tyrosine residue that has not been
 CC modified by phosphate or similar charged group. The non-phosphorylated
 CC peptide is used to inhibit a signal transduction process that involves
 CC binding of a phosphorylated protein or peptide to the SH2 domain of a
 CC signal transduction protein, particularly Grb2. It is used specifically
 CC for treatment of hyper-proliferative diseases, especially human cancer
 XX Sequence 9 AA;
 SQ

Query Match 86.0%; Score 37; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSYVNVQ 7
 DB 3 PSYVNVQ 9

RESULT 13
 AAU10467
 ID AAU10467 standard; peptide; 9 AA.
 AC AAU10467;
 XX
 XX 14-FEB-2002 (first entry)
 DT
 XX
 DE Nucleic acid tagged cancer related peptide AS15.
 XX
 KW Tagged array; cancer; cytostatic; multifactorial disease; PNA;
 KW peptide-nucleic acid; AS15.
 XX
 OS Homo sapiens.
 XX US2001031469-A1.
 PN 18-OCT-2001.
 XX
 PD 02-JAN-2001; 2001US-00753114.
 PF
 XX 03-JAN-2000; 2000US-0174171P.
 PR (VOLI/) VOLINIA S.
 XX
 PA Volinia S;
 PI
 XX WPI; 2002-025218/03.
 DR
 XX

Analysis of protein or enzyme activity in samples, e.g., biopsies,
 comprises using a pool of tagged substrates which can then be sorted onto
 a solid surface array after reaction.

Example 3; Page 7; 36pp; English.

The invention relates to analysing the activity or level of one or more
 proteins or enzymes, comprising providing a pool of substrates (e.g.
 peptides, antibodies, binding domains or other molecules which act as
 substrates or control substrates), each with a specific tag and
 representing a substrate of one or more of the proteins or enzymes, or
 substrates derived from these using the tagged substrates as substrates,
 hybridising the pool of tagged substrates to an ordered array of specific
 and complementary tags immobilised on a surface, where the array
 comprises different tags, at least some of which are control tags, each
 tag is localised in a predetermined region of the surface and the density
 of different tags is greater than 100 different tags per square
 centimetre, and all tags in the substrates derived using the proteins or
 enzymes are complementary to at least some of the immobilised tags,
 quantifying hybridisation of the substrates tagged with nucleic acids or
 peptide nucleic acids (PNAs) to the array, where the quantification is
 proportional to the activity of proteins or enzymes which modify or
 attach to the substrates tagged with nucleic acids or PNAs. The method
 can be used for detection of post-translationally modified proteins and
 for identifying target proteins capable of binding to, or serving as,

enzymes or molecular adapters involved in biological functions. They can be used, e.g., for analysing biopsies from cancers and other multifactorial diseases. They can be used to identify previously unknown proteins or new substrates. The use of tagged substrates rather than immobilised substrates leads to increased stability of the substrate, improved quality control and lower production costs. The tagged substrate can be kept lyophilised until use, separate from other tagged substrates. Substrates can be changed, refined or differentially labelled at any time, without the need for designing or printing a new tag array. The present sequence is a hybridised array sequence which is attached to a solid support via its 5' amide modification. The sequence a target peptide-nucleic acid sequence (PNA) which comprises a peptide derived from a human cancer related protein and a nucleic acid sequence complementary to the bound nucleic acid sequences presented as AAS16066-AAS16083 and is used to demonstrate the method of the invention

Sequence 9 AA;

Query Match 86.0%; Score 37; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 3 PSYVNVQ 9

RESULT 14

AAU10458
ID AAU10458 standard; peptide; 9 AA.

AC AAU10458;

DT 14-FEB-2002 (first entry)

DE Nucleic acid tagged cancer related peptide AS6.

KW Tagged array; cancer; cytostatic; multifactorial disease; PNA;
KW peptide-nucleic acid; AS6.

OS Homo sapiens.

PN US2001031469-A1.

PD 18-OCT-2001.

PF 02-JAN-2001; 2001US-00753114.

PR 03-JAN-2000; 2000US-0174171P.

PA (VOLI/) VOLINIA S.

PI Volinia S;

DR WPI; 2002-025218/03.

PT Analysis of protein or enzyme activity in samples, e.g., biopsies,
PT comprises using a pool of tagged substrates which can then be sorted onto
PT a solid surface array after reaction.

PS Example 3; Page 7; 36pp; English.

XX The invention relates to analysing the activity or level of one or more
XX proteins or enzymes, comprising providing a pool of substrates (e.g.
XX peptides, antibodies, binding domains or other molecules which act as
XX substrates or control substrates), each with a specific tag and
XX representing a substrate of one or more of the proteins or enzymes, or
XX substrates derived from these using the tagged substrates as substrates,
XX hybridising the pool of tagged substrates to an ordered array of specific
XX and complementary tags immobilised on a surface, where the array
XX comprises different tags, at least some of which are control tags, each
XX tag is localised in a predetermined region of the surface and the density

of different tags is greater than 100 different tags per square centimetre, and all tags in the substrates derived using the proteins or enzymes are complementary to at least some of the immobilised tags, quantifying hybridisation of the substrates tagged with nucleic acids or peptide nucleic acids (PNAs) to the array, where the quantification is proportional to the activity of proteins or enzymes which modify or attach to the substrates tagged with nucleic acids or PNAs. The method can be used for detection of post-translationally modified proteins and for identifying target proteins capable of binding to, or serving as, enzymes or molecular adapters involved in biological functions. They can be used, e.g., for analysing biopsies from cancers and other multifactorial diseases. They can be used to identify previously unknown proteins or new substrates. The use of tagged substrates rather than immobilised substrates leads to increased stability of the substrate, improved quality control and lower production costs. The tagged substrate can be kept lyophilised until use, separate from other tagged substrates. The quality of each tagged compound can be verified at any stage. Substrates can be changed, refined or differentially labelled at any time, without the need for designing or printing a new tag array. The present sequence is a hybridised array sequence which is attached to a solid support via its 5' amide modification. The sequence a target peptide-nucleic acid sequence (PNA) which comprises a peptide derived from a human cancer related protein and a nucleic acid sequence complementary to the bound nucleic acid sequences presented as AAS16066-AAS16083 and is used to demonstrate the method of the invention

Sequence 9 AA;

Query Match 86.0%; Score 37; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 3 PSYVNVQ 9

RESULT 15

AAE15236
ID AAE15236 standard; peptide; 9 AA.

AC AAE15236;

DT 07-MAR-2002 (first entry)

DE SHC phospho peptide, used to determine pTyr mimetics.

KW N-oxalyl peptide; therapeutic; prophylactic; treatment; psoriasis;
KW downstream signal transduction pathway; tumour.

OS Unidentified.

PH Key Location/Qualifiers

FT Modified-site 5 /note= "Phosphorylated tyrosine"

XX US6307090-B1.

XX 23-OCT-2001.

XX 22-JAN-1999; 99US-00236160.

XX 22-JAN-1999; 99US-00236160.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Burke TR, Yao Z, King CR;

XX WPI; 2002-033269/04.

XX New acylated oligopeptides having cell signal inhibiting activity are
XX useful in the treatment of e.g. psoriasis.

PS Example; Fig 1; 42pp; English.

XX The present invention relates to pharmaceutically active compounds
CC comprising N-oxalyl peptide structure. The compounds of the invention are
CC used in the therapeutic or prophylactic treatment of warm-blooded animal
CC or human body e.g. tumour and psoriasis. They are useful in the
CC diagnostic treatment of diseases depending on the downstream signal
CC transduction pathway. The present sequence is SHC phospho peptide which
CC is used to determine pTyr mimetics in the context of Grb2 SH2 inhibitors.
CC This peptide is used in the exemplification of the invention

XX Sequence 9 AA;

Query Match 86.0%; Score 37; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 3 PSYVNVQ 9
|||||

RESULT 16

ABP33784

ID ABP33784 standard; protein; 61 AA.

XX ABP33784;

XX 08-JUL-2002 (first entry)

DE Human ORF2757 protein, SEQ ID NO:5514.

XX Human; ORF: open reading frame; ORFX: drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.

XX Homo sapiens.

XX WO200190366-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US017076.

XX 24-MAY-2000; 2000US-0206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shimkets RA;

XX WPI; 2002-106200/14.

XX N-PSDB; ABN77810.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation.

XX Claim 10; Page 1629; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins

CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-

CC ABN79587 represent cDNAs encoding them. The invention also encompasses

CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antinefractive activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases

XX Sequence 61 AA;

Query Match 83.7%; Score 36; DB 5; Length 61;
Best Local Similarity 85.7%; Pred. No. 5.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 54 PSYVNIQ 60
|||||

RESULT 17

AAW78664

ID AAW78664 standard; peptide; 8 AA.

XX AAW78664;

XX 04-NOV-1998 (first entry)

XX SH2 domain binding inhibiting peptide SEQ ID NO:157.

XX SH2 domain; binding; inhibition; interaction; site specific;
KW signal transduction; protein tyrosine kinase; phosphotyrosine;
KW growth factor receptor; oncogene; cellular growth; cell proliferation;
KW metabolic control; diabetes; FTK; proto-oncogene; insulin receptor.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 3 /note= "unspecified"

XX US5801149-A.

XX 01-SEP-1998.

XX 21-MAR-1995; 95US-00408604.

XX 19-JUN-1991; 91US-00722359.

XX 09-OCT-1992; 92US-00959949.

XX 08-OCT-1993; 93US-00134558.

XX PA (JOSL-) JOSLIN DIABETES CENT INC.
XX PI Shoelson S;
XX DR WPI; 1998-494822/42.
XX DR WPI; 1998-494822/42.
XX PT Inhibiting site-specific SH2 domain interaction - with peptide containing
XX PT phosphotyrosine or phosphotyrosine mimic.
XX PS Disclosure; Col 99; 70pp; English.
XX CC A method has been developed of inhibiting a site-specific interaction
XX CC between a first molecule having an SH2 domain and a second molecule that
XX CC interacts with the SH2 domain. The method comprises contacting the first
XX CC molecule with a 4- to 30-mer peptide containing a sequence of formula: R1
XX CC -Met-R3-Met (I), where R1 = phosphotyrosine (pTyr) or a phosphotyrosine
XX CC analogue having a hydrolysis-resistant phosphorous moiety, and R3 = any
XX CC amino acid. AAW78501 to AAW78523 represent specifically claimed examples
XX CC of the peptides described. The peptides are useful for modulating both
XX CC cellular growth to control unwanted cell proliferation in e.g. selected
XX CC malignancies and for metabolic control in e.g. diabetes, by inhibiting
XX CC signal transduction molecules such as protein tyrosine kinases (PTKs)
XX CC which include growth factor receptors, proto-oncogene and oncogene
XX CC products and the insulin receptor. The peptides are also useful for
XX CC treating and for studying the enzymatic mechanisms of PTPase activity and
XX CC the metabolic and biochemical roles of PTPases. AAW78524 to AAW78702
XX CC represent other peptides given in the present invention, but which are
XX CC not specifically claimed
XX SQ Sequence 8 AA;
XX
Query Match 81.4%; Score 35; DB 2; Length 8;
Best Local Similarity 87.5%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVNVQN 8
DB 1 PSXVNVQN 8
RESULT 18
AAW78663
ID AAW78663 standard; peptide; 9 AA.
XX AC AAW78663;
XX DT 04-NOV-1998 (first entry)
XX DE SH2 domain binding inhibiting peptide SEQ ID NO:156.
XX KW SH2 domain; binding; inhibition; interaction; site specific;
XX KW signal transduction; protein tyrosine kinase; phosphotyrosine;
XX KW growth factor receptor; oncogene; cellular growth; cell proliferation;
XX KW metabolic control; diabetes; PTK; proto-oncogene; insulin receptor.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 3 /note= "unspecified"
FT FT
FT FT
XX PN US5801149-A.
XX XX
XX PD 01-SEP-1998.
XX XX
XX PF 21-MAR-1995; 95US-00408604.
XX PR 19-JUN-1991; 91US-00722359.
XX PR 09-OCT-1992; 92US-00959949.
XX PR 08-OCT-1993; 93US-00134558.
XX PA (JOSL-) JOSLIN DIABETES CENT INC.
XX PI Shoelson S;

XX PI Shoelson S;
XX DR WPI; 1998-494822/42.
XX DR WPI; 1998-494822/42.
XX PT Inhibiting site-specific SH2 domain interaction - with peptide containing
XX PT phosphotyrosine or phosphotyrosine mimic.
XX PS Disclosure; Col 99; 70pp; English.
XX CC A method has been developed of inhibiting a site-specific interaction
XX CC between a first molecule having an SH2 domain and a second molecule that
XX CC interacts with the SH2 domain. The method comprises contacting the first
XX CC molecule with a 4- to 30-mer peptide containing a sequence of formula: R1
XX CC -Met-R3-Met (I), where R1 = phosphotyrosine (pTyr) or a phosphotyrosine
XX CC analogue having a hydrolysis-resistant phosphorous moiety, and R3 = any
XX CC amino acid. AAW78501 to AAW78523 represent specifically claimed examples
XX CC of the peptides described. The peptides are useful for modulating both
XX CC cellular growth to control unwanted cell proliferation in e.g. selected
XX CC malignancies and for metabolic control in e.g. diabetes, by inhibiting
XX CC signal transduction molecules such as protein tyrosine kinases (PTKs)
XX CC which include growth factor receptors, proto-oncogene and oncogene
XX CC products and the insulin receptor. The peptides are also useful for
XX CC treating and for studying the enzymatic mechanisms of PTPase activity and
XX CC the metabolic and biochemical roles of PTPases. AAW78524 to AAW78702
XX CC represent other peptides given in the present invention, but which are
XX CC not specifically claimed
XX SQ Sequence 9 AA;
XX
Query Match 81.4%; Score 35; DB 2; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.4e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVNVQN 8
DB 1 PSXVNVQN 8
RESULT 19
AAW78661
ID AAW78661 standard; peptide; 16 AA.
XX AC AAW78661;
XX DT 04-NOV-1998 (first entry)
XX DE SH2 domain binding inhibiting peptide SEQ ID NO:154.
XX KW SH2 domain; binding; inhibition; interaction; site specific;
XX KW signal transduction; protein tyrosine kinase; phosphotyrosine;
XX KW growth factor receptor; oncogene; cellular growth; cell proliferation;
XX KW metabolic control; diabetes; PTK; proto-oncogene; insulin receptor.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 8 /note= "unspecified"
FT FT
FT FT
XX PN US5801149-A.
XX XX
XX PD 01-SEP-1998.
XX XX
XX PF 21-MAR-1995; 95US-00408604.
XX PR 19-JUN-1991; 91US-00722359.
XX PR 09-OCT-1992; 92US-00959949.
XX PR 08-OCT-1993; 93US-00134558.
XX PA (JOSL-) JOSLIN DIABETES CENT INC.
XX PI Shoelson S;

Query Match 79.1%; Score 34; DB 6; Length 187;
 Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYNNVQ 7
 DB 81 PSYNNVE 87

RESULT 22
 ADC94754
 ID ADC94754 standard; protein; 251 AA.

XX AC ADC94754;
 XX DT 01-JAN-2004 (first entry)
 XX DE E. faecium protein sequence SEQ ID 4381.
 XX KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 XX KW abdominal-pelvic infection.

XX OS Enterococcus faecium.

XX PN US6583275-B1.

XX PD 24-JUN-2003.

XX PF 30-JUN-1998; 98US-00107532.

XX PR 02-JUL-1997; 97US-0051571P.

XX PR 14-MAY-1998; 98US-0085598P.

XX FA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI; 2003-799836/75.

XX DR N-PSDB; ADC91100.

XX PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 Enterococcus faecium polypeptide useful for detection, prevention and
 treatment of a pathological condition resulting from a bacterial
 infection.

XX PS Example 1; SEQ ID NO 4381; 243pp; English.

XX CC The invention relates to an isolated nucleic acid derived from
 Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 one of 10 fully defined sequences given in the (or comprising 40
 sequential nucleotides chosen from any of the nucleic acids, its
 complement or sequences hybridising to it). Also included are a
 recombinant vector comprising the nucleic acid operably linked to
 transcription regulatory element, a cell comprising the vector and a
 single-stranded probe comprising the nucleic acid. The nucleic acids are
 chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 XX CC The nucleic acids are useful for diagnosing pathological conditions
 resulting from E. faecium bacterial infection (e.g. urinary tract
 infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 infection) and for screening drugs such as agonists and antagonists. The
 nucleic acid is useful for recombinant production of Candida albicans -
 derived peptides or antisense polypeptides. Pharmaceutical compositions
 XX CC and vaccines containing the nucleic acid are useful for preventing or
 treating Enterococcus faecium infections. The present sequence represents
 one if the disclosed E. faecium proteins.

XX SQ Sequence 251 AA;

Query Match 79.1%; Score 34; DB 7; Length 251;
 Best Local Similarity 85.7%; Pred. No. 72;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYNNVQ 7
 DB 222 PAYNNVQ 228

RESULT 23
 ADA34556

ID ADA34556 standard; protein; 273 AA.

XX AC ADA34556;

XX DT 20-NOV-2003 (first entry)

XX DE Acinetobacter baumannii protein #1717.

XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX PN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX FA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX DR N-PSDB; ADA30430.

XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 for diagnosing a bacterial disease, as components of antibacterial
 vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 plants.

XX PS Example; SEQ ID NO 5843; 328pp; English.

XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 The A. baumannii nucleic acids and polypeptides are useful as reagents
 for diagnosing a bacterial disease, as components of antibacterial
 vaccines, as targets for antibacterial drugs, to detect the presence of
 A. baumannii and other Acinetobacter species in a sample, in screening
 compounds for the ability to interfere with the A. baumannii life cycle
 or to inhibit A. baumannii infection, and as biocontrol agents for
 plants. The present sequence represents the amino acid sequence of an A.
 baumannii protein.

XX SQ Sequence 273 AA;

Query Match 79.1%; Score 34; DB 6; Length 273;
 Best Local Similarity 71.4%; Pred. No. 79;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYNNVQN 8
 DB 125 SYNNIQN 131

RESULT 24
 ABU24418

ID ABU24418 standard; protein; 471 AA.

XX AC ABU24418;

XX DT 19-JUN-2003 (first entry)

XX DB Protein encoded by Prokaryotic essential gene #9945.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Clostridium botulinum.
OS WO20027183-A2.
PN XX
PD 03-OCT-2002.
PF XX
PP 21-MAR-2002; 2002WO-US009107.
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
PI WPI: 2003-029926/02.
XX N-PSDB; ACR28288.
DR
DR New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 52342; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 471 AA;

Query Match 79.1%; Score 34; DB 6; Length 471;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVNVON 8
DB 114 PSYVNLPN 121

RESULT 25
AAM95180
ID AAM95180 standard; protein; 144 AA.
XX
AC AAM95180;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 3838.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
OS Homo sapiens.
XX
FN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 30-JUN-2000; 2000US-0215135P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0218880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.

R 07-JUN-2000; 2000US-0209467P.
R 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer.

Claim 11; SEQ ID NO 1264; 766pp; English.

The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system, reproductive system, immune, respiratory, neurological and gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a protein of the invention

XX

SQ Sequence 144 AA;
 Query Match 76.7%; Score 33; DB 4; Length 144;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PSYVNVQ 7
 |||:|:
 Db 51 PSYINLQ 57

RESULT 27
 AAB92664
 ID AAB92664 standard; protein; 215 AA.
 XX
 AC AAB92664;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11022.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX
 PS Claim 8; SEQ ID NO 11022; 2537pp + Sequence Listing; English.
 XX

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent polynucleotides, all of which are used in the exemplification of the present invention

XX
 SQ Sequence 215 AA;
 Query Match 76.7%; Score 33; DB 4; Length 215;
 Best Local Similarity 85.7%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 PSYVNVQ 7
 |||:|:
 Db 69 PGYVNVQ 75

RESULT 28
 AAB88321
 ID AAB88321 standard; protein; 215 AA.
 XX
 AC AAB88321;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Human membrane or secretory protein clone PSEC0008.
 XX
 KW Human; secretory protein; membrane protein; vaccine; gene therapy; rheumatoid arthritis; diabetes.
 XX
 OS Homo sapiens.
 XX
 PN EPI067182-A2.
 XX
 PD 10-JAN-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114090.
 XX
 PR 08-JUL-1999; 99JP-00194179.
 PR 11-JAN-2000; 2000JP-00118775.
 PR 02-MAY-2000; 2000JP-00183766.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 PI WPI; 2001-093989/11.
 XX
 DR N-PSDB; AAF933748.
 XX
 PT Nucleic acids encoding secretory proteins/membrane proteins, useful in gene therapy or as candidate target molecules in drug development.
 XX
 PS Claim 1; SEQ ID NO 10; 609pp + Sequence Listing; English.
 XX

This invention relates to nucleic acid sequences AAF93744 - AAF93916 which encode human secretory or membrane proteins represented by AAB88317 - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the invention. The invention also includes methods for the production of antibodies directed against the proteins, and cDNA sequences, which can be used in vaccines. The polynucleotide sequences can be used in gene therapy. The polynucleotide sequences and the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate secretory protein/membrane protein expression. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples. They may also be used to study the expression and function of secretory proteins/membrane polypeptides and their role in metabolism. The polypeptides may be used as antigens in the production of antibodies against them and in assays to identify modulators (agonists and antagonists) of expression and activity. The antibodies and antagonists may also be used as therapeutic agents to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the polypeptides in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Examples of diseases which may be treated include rheumatoid arthritis and diabetes

SQ Sequence 215 AA;
 Query Match 76.7%; Score 33; DB 4; Length 215;
 Best Local Similarity 85.7%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 69 PGYVNVQ 75
 | | | | |
 Db 69 PGYVNVQ 75

RESULT 29
 ADB65101
 ID ADB65101 standard; protein; 256 AA.
 XX
 AC ADB65101;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human protein encoded by clone SPLEN20015100.
 XX
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 PN EP1308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 PF 28-MAR-2002; 2002EP-00007401.
 XX
 PR 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX

Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 WPI; 2003-450961/43.
 DR N-PSDB; ADB63131.
 XX

New polynucleotides and polypeptides, useful for developing a diagnostic
 marker or medicines for regulation of their expression and activity, or
 as targets of gene therapy.
 Claim 1; Page; 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected
 from 1970 fully defined nucleotide sequences which encode novel
 polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 or its partial peptide, an antibody binding to the polypeptide or peptide
 of the polynucleotide, immunologically assaying the polypeptide or
 peptide of the polynucleotide by contacting the polypeptide or peptide
 with the antibody of the encoded protein, and observing the binding
 between the two, a transformant carrying the polynucleotide in an
 expressible manner and an antisense polynucleotide. The oligonucleotide
 is useful as a primer for synthesising the polynucleotide, or as a probe
 for detecting the polynucleotide. The polynucleotides and encoded
 proteins are useful as pharmaceutical agents and many disease-related
 genes may be included in them, for developing a diagnostic marker or
 medicines for regulation of their expression and activity, or as targets
 of gene therapy. The genes are involved in tissue and/or cell
 regeneration. Membrane proteins, signal transduction-related proteins,
 transcription-related proteins, disease-related proteins and genes
 encoding them can be used as indicators for diseases (e.g. osteoporosis,
 neurological diseases, cancer, tumours). The cDNA may be used to regulate
 the activity or expression of the encoded protein to treat diseases. The

CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 256 AA;
 Query Match 76.7%; Score 33; DB 7; Length 256;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 98 PGYVNVQ 104
 | | | | |
 Db 98 PGYVNVQ 104

RESULT 30
 AAY94890
 ID AAY94890 standard; protein; 267 AA.
 XX
 AC AAY94890;
 XX
 DT 12-JUN-2000 (first entry)
 XX
 DE Human protein clone HP02798.
 XX
 KW Human protein; hydrophobic domain; nutritional source; haematopoiesis;
 KW cytokine production; cell proliferation; cell differentiation;
 KW immune deficiency; infectious disease; autoimmune disorder; asthma;
 KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
 KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
 KW nervous system disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
 KW systemic cytokine damage; tissue differentiation; contraceptive injury;
 KW coagulation disorder; myocardial infarction; inflammatory condition;
 KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
 KW nephritis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200005367-A2.
 XX
 PD 03-FEB-2000.
 XX
 PF 22-JUL-1999; 99WO-JP003929.
 XX
 PR 24-JUL-1998; 98JP-00208820.
 PR 07-AUG-1998; 98JP-00224105.
 PR 25-AUG-1998; 98JP-00238116.
 PR 09-SEP-1998; 98JP-00254736.
 PR 29-SEP-1998; 98JP-00275505.
 XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI; 2000-182694/16.
 XX

Novel human proteins having hydrophobic domains useful for treating
 osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple
 sclerosis, rheumatoid arthritis, cancer, anemia, and stroke.
 Claim 1; Page 316-317; 351pp; English.

This sequence represents a human protein of the invention, which has
 hydrophobic domains. The DNA sequences can be used as a probe or as a
 genetic marker. The protein can also be used as a marker, and to identify
 potential genetic disorders. The DNA and protein can also be used as
 nutritional sources or supplements. The protein exhibits cytokine, cell
 proliferation, cell differentiation activities and induces production of
 other cytokines in certain cell populations. The protein also exhibits
 immune stimulating or immune suppressing activity. It can be used in the

CC treatment of various immune deficiencies and disorders, and to treat
 CC infectious diseases caused by viral, bacterial, fungal or other
 CC infections. The protein is also used for treating autoimmune disorders
 CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
 CC arthritis. It is also useful in the treatment of allergic reactions and
 CC conditions such as asthma, and in immune suppression after organ
 CC transplantation. The protein is useful in regulation of haematopoiesis
 CC and consequently in the treatment of myeloid or lymphoid cell
 CC deficiencies. It is also used in compositions for tissue growth or
 CC regeneration. The protein is also used in the treatment of osteoporosis
 CC or osteoarthritis and in the treatment of periodontal disease and other
 CC tooth repair processes. The protein is used in the treatment of nervous
 CC system disorders such as Alzheimer's disease, Parkinson's disease, and
 CC Huntington's disease. They are useful for protection or regeneration and
 CC treatment of lung or liver fibrosis, reperfusion injury in various
 CC tissues, and conditions resulting from systemic cytokine damage. They are
 CC also used for promoting or inhibiting tissue differentiation. They are
 CC also used as contraceptives since they exhibit activin or inhibin related
 CC activities and as a fertility inducing therapeutic. They are used for
 CC treating various coagulation disorders and in treatment and prevention of
 CC conditions resulting from coagulation activities e.g. myocardial
 CC infarction or stroke. They also acts as receptors, receptor ligands or
 CC inhibitors or agonists of receptor/ligand interactions. They are used to
 CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
 CC reperfusion injury, arthritis, and nephritis. They can be used to prevent
 CC tumours

XX SQ Sequence 267 AA;

Query Match 76.7%; Score 33; DB 3; Length 267;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNVQ 7
 | | | | |
 Db 69 PGYVNVQ 75

RESULT 31
 AAB68874
 ID AAB68874 standard; protein; 267 AA.

XX AC AAB68874;
 XX DT 24-APR-2001 (first entry)
 XX DE Human RECAP polypeptide, SEQ ID NO: 4.
 XX KW Human; RECAP; receptors and associated proteins; cerebroprotective;
 KW neurotropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
 KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
 KW antithyroid; immunosuppressive; nephrotropic; antitumor; thyromimetic;
 KW cytostatic; antibacterial; virucide; fungicide; protozoacide;
 KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer.

XX OS Homo sapiens.

XX PN WO200107612-A2.

XX PD 01-FEB-2001.

XX PF 21-JUL-2000; 2000WO-US020035.

XX PR 21-JUL-1999; 99US-0145232P.

XX PR 07-OCT-1999; 99US-0158578P.

XX PR 12-NOV-1999; 99US-0165192P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;

XX PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;

XX DR WPI; 2001-168554/17.

DR N-PSDB; AAF58598.

XX Novel receptors and associated proteins for diagnosis and treatment of
 PT neurological disorders, immunological disorders including autoimmune/
 PT inflammatory disorders and cell proliferative disorders such as cancer.
 XX Claim 1; Page 96-97; 128pp; English.

XX The present sequence is a human RECAP (receptors and associated proteins)
 CC polypeptide. RECAP polynucleotides and polypeptides are useful in the
 CC diagnosis, treatment and prevention of neurological disorders such as
 CC stroke, Alzheimer's disease, Pick's disease, Huntington's disease,
 CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral
 CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD
 CC (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
 CC syndrome); immunological disorders, including autoimmune/inflammatory
 CC disorders such as AIDS, DiGeorge's syndrome, severe combined
 CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
 CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
 CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
 CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
 CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
 CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis and cancer

XX SQ Sequence 267 AA;

Query Match 76.7%; Score 33; DB 4; Length 267;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNVQ 7
 | | | | |
 Db 69 PGYVNVQ 75

RESULT 32
 AAB65010
 ID AAB65010 standard; protein; 267 AA.

XX AC AAB65010;

XX DT 23-MAR-2001 (first entry)

XX DE Human secreted protein #18.

XX KW Secreted protein; gene therapy; vaccine; cancer; leukemia;
 KW autoimmune disease; allergy; inflammation; graft rejection;
 KW hyperproliferation; cardiovascular; infection.

XX OS Homo sapiens.

XX PN WO200075375-A1.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-US015187.

XX PR 07-JUN-1999; 99US-0137725P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;

XX PI Laflleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PE;

XX DR WPI; 2001-061741/07.

XX PT Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
 PT preventing, diagnosing and/or treating cancers and for promoting wound
 PT healing.

XX PS Claim 1; Page 468-469; 530pp; English.

CC The present invention relates to 26 secreted human proteins. The proteins
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. For example, they
 CC may be used in gene therapy or in vaccines. Typical of diseases which are
 CC potentially treatable are cancers (including leukemia), autoimmune
 CC diseases, allergies, inflammation, graft rejection, hyperproliferation,
 CC cardiovascular diseases (particularly critical limb ischemia and coronary
 CC disease) and any involving abnormal angiogenesis, neurodegeneration
 CC and/or infectious diseases
 CC
 XX Sequence 267 AA;

Query Match 76.7%; Score 33; DB 4; Length 267;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 | | | | |
 Db 69 PGYVNVQ 75

RESULT 33
 ABP69845
 ID ABP69845 standard; protein; 322 AA.

AC ABP69845;

DT 20-JAN-2003 (first entry)

DE Human polypeptide SEQ ID NO 1892.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI: 2002-759812/82.

XX N-PSDB; ABZ12062.

XX New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.

XX Claim 9; SEQ ID NO 1892; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
 CC ABZ12066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP69802-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical

CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 322 AA;

Query Match 76.7%; Score 33; DB 5; Length 322;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 | | | | |
 Db 124 PGYVNVQ 130

RESULT 34

ABM73375

ID ABM73375 standard; protein; 577 AA.

AC ABM73375;

DT 20-NOV-2003 (first entry)

DE Staphylococcus aureus protein #2615.

XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 KW enzymatic assay; antibiotic target.

XX Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-IB002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Massignani V, Mora M, Scarselli M;

XX WPI: 2003-120786/11.

XX N-PSDB; ACF74935.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or
 PT preventing Staphylococcal infection, specifically an infection caused by
 PT S. aureus, e.g. sepsis.

XX Claim 1; SEQ ID NO 5230; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention

XX Sequence 577 AA;

Query Match 76.7%; Score 33; DB 6; Length 577;
 Best Local Similarity 75.0%; Pred. No. 3e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 DE |||||
 Db 188 PAYHNVQN 195
 |||||

RESULT 35
 ABB71917
 ID ABB71917 standard; protein; 732 AA.
 XX AC ABB71917;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 42543.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL16020.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 42543; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 732 AA;
 Query Match 76.7%; Score 33; DB 4; Length 732;
 Best Local Similarity 62.5%; Pred. No. 4e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 DE |||||
 Db 259 PSYLVNMRN 266
 |||||

RESULT 36
 AAY53621
 ID AAY53621 standard; protein; 880 AA.
 XX AC AAY53621;
 XX DT 12-SEP-2003 (revised)
 XX DT 22-FEB-2000 (first entry)

XX Amino acid sequence of beta-galactosidase encoded by the lactose operon.
 DE
 XX Lactose operon; permease; beta-galactosidase; mutant; yoghurt;
 KW beta-galactosidase activity; lactose fermentation;
 KW fermented milk product; Streptococcus thermophilus.
 XX OS Lactobacillus delbrueckii subsp. bulgaricus.
 XX FR2778921-A1.
 XX PN 26-NOV-1999.
 XX PD 22-MAY-1998; 98FR-00006456.
 XX PF 22-MAY-1998; 98FR-00006456.
 XX PR (DANO-) CIE DANONE SA GERVAIS.
 XX PA Benbadis L, Brignon P, Gendre F;
 XX WPI; 2000-026146/03.
 XX N-PSDB; AAZ36227.
 XX New mutant Lactobacillus bulgaricus strain lacking beta-galactosidase
 PT activity useful for making fermented milk products, especially yogurt.
 XX Disclosure; Page 22-27; 32pp; French.

CC AAY53620-21 are encoded by the lactose operon of Lactobacillus
 CC bulgaricus, and represent a permease (AAY53620) and beta-galactosidase
 CC (AAY53621). The specification describes a mutant L. bulgaricus strain
 CC lacking beta-galactosidase activity, because it has a termination
 CC (nonsense) mutation in at least one of the coding sequences of the
 CC lactose operon. The resulting mutant strain is unable to ferment lactose.
 CC Since the mutant strain cannot ferment lactose, fermented milk products
 CC produced using it do not have to be cooled to prevent post-acidification.
 CC The mutant L. bulgaricus strain of the invention is used for producing a
 CC fermented milk product, especially yoghurt. Especially, milk is fermented
 CC with a culture of the mutant strain and optionally Streptococcus
 CC thermophilus in the presence of at least one sugar assimilable by the
 CC mutant strain, preferably glucose. (Updated on 12-SEP-2003 to standardise
 CC OS field)

XX SQ Sequence 880 AA;
 Query Match 76.7%; Score 33; DB 3; Length 880;
 Best Local Similarity 85.7%; Pred. No. 4.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 DE |||||
 Db 101 PQYVNVQ 107
 |||||

RESULT 37
 ABB66417
 ID ABB66417 standard; protein; 4345 AA.
 XX AC ABB66417;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 26043.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.


```

RESULT 40
ABU03776
ID ABJ03776 standard; protein; 82 AA.
XX
AC ABJ03776;
XX
DT 25-SEP-2002 (first entry)
XX
DE Human ovary specific protein SEQ ID NO: 218.
XX
KW Human; ovary; ovarian cancer; ovarian disease; gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200240720-A2.
XX
PD 23-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-US045010.
XX
PR 20-NOV-2000; 2000US-0249997P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Recipon H, Cafferkey R, Sun Y, Liu C;
XX
DR WPI; 2002-547588/58.
XX
XX New ovary polypeptides useful for detecting, diagnosing, monitoring,
PT treating, staging and imaging cancers in humans having cancer and non-
PT cancerous ovary disease.
XX
PS Claim 11; Page 289-290; 296pp; English.
XX
CC The present invention provides human proteins and coding sequences
CC specifically found in ovary cells. These can be used in the diagnosis and
CC treatment of ovarian diseases, including cancer. The present sequence is
CC a protein of the invention
XX
SQ Sequence 82 AA;
Query Match 74.4%; Score 32; DB 5; Length 82;
Best Local Similarity 62.5%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVNVQN 8
Db |||||:|
13 PSYINKN 20

RESULT 41
AAU73415
ID AAU73415 standard; protein; 141 AA.
XX
AC AAU73415;
XX
DT 29-FEB-2000 (first entry)
XX
DE Human secreted protein clone yb124_1 protein sequence SEQ ID NO:52.
XX
KW Human; secreted protein; immunostimulatory; haemostatic; cytokine;
KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
KW thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO9958642-A2.
XX
PD 18-NOV-1999.
XX

```

```

PF 14-MAY-1999; 99WO-US010843.
XX
PR 14-MAY-1998; 98US-0085472P.
PR 17-AUG-1998; 98US-0096824P.
PR 11-SEP-1998; 98US-0099843P.
PR 11-SEP-1998; 98US-0099950P.
PR 15-SEP-1998; 98US-0100424P.
PR 29-SEP-1998; 98US-0102329P.
PR 09-OCT-1998; 98US-0103615P.
PR 11-DEC-1998; 98US-0111799P.
PR 14-DEC-1998; 98US-0112159P.
PR 31-DEC-1998; 98US-0114415P.
PR 10-FEB-1999; 98US-00248059.
PR 06-APR-1999; 99US-00287150.
PR 13-MAY-1999; 99US-00311021.
XX
PA (GEMY) GENETICS INST INC.
XX
XX Wong GG, Clark HF, Fechtel K, Agostino MJ;
PI
XX
DR WPI; 2000-053095/04.
DR N-FSDB; AAZ52500.
XX
XX Novel polynucleotides and proteins having biological activities which
PT make them suitable for treating, preventing or ameliorating medical
PT conditions in humans or animals.
XX
PS Claim 61; Page 618; 730pp; English.
XX
XX The present invention describes human secreted proteins encoded by
CC polynucleotides obtained from adult testes, foetal brain, adult brain,
CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
CC cDNA libraries. The polynucleotides and proteins are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals.
CC Suggested activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
CC suppressor activity, and tumour inhibition activity. The polynucleotides
CC are also stated to be useful for gene therapy. Therapeutic compositions
CC are also presently valuable for veterinary applications. AAZ52475 to
CC AAZ52581 encode human secreted proteins, and AAU73390 to AAU73500
CC represent human secreted proteins, given in the present invention
XX
SQ Sequence 141 AA;
Query Match 74.4%; Score 32; DB 3; Length 141;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYVNVQN 8
Db |||||:|
37 PSYFNQON 44

RESULT 42
AAU23734
ID AAU23734 standard; protein; 200 AA.
XX
AC AAU23734;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #820.
XX
KW Human; oxidoreductase enzyme; transferase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW

```

KW nephrotropic; anticoagulant.

XX Homo sapiens.

XX W0200155301-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001239.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209457P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216980P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225470P.

PR 14-AUG-2000; 2000US-0225577P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0235759P.

PR 22-AUG-2000; 2000US-0236279P.

PR 22-AUG-2000; 2000US-0236681P.

PR 22-AUG-2000; 2000US-0236688P.

PR 23-AUG-2000; 2000US-0237182P.

PR 30-AUG-2000; 2000US-0227009P.

PR 01-SEP-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234224P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465566/50.
 XX N-PSDB; AAS41604.
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
 PT treating neural, immune system, muscular, reproductive, pulmonary,
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.
 XX Claim 11; SEQ ID NO 1730; 1180pp; English.
 XX The present invention relates to the isolation of novel human enzyme
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
 CC infectious disorders (e.g. Influenza). The polynucleotides of the
 CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
 CC the novel human enzyme polypeptides of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 200 AA;

Query Match 74.4%; Score 32; DB 4; Length 200;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 Db 111 PDYVNIQ 117

RESULT 43
 AAU23196
 ID AAU23196 standard; protein; 200 AA.
 AC AAU23196;
 DT 17-DEC-2001 (first entry)
 DE Novel human enzyme polypeptide #282.
 XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.
 XX Homo sapiens.
 OS
 XX WO20015301-A2.
 PN
 XX 02-AUG-2001.
 PD
 XX
 FF 17-JAN-2001; 2001WO-US001239.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR

16-MAR-2000; 2000US-0189874P.
 17-MAR-2000; 2000US-0190076P.
 18-APR-2000; 2000US-0198123P.
 19-MAY-2000; 2000US-0205515P.
 07-JUN-2000; 2000US-0209467P.
 28-JUN-2000; 2000US-0214886P.
 30-JUN-2000; 2000US-0215135P.
 07-JUL-2000; 2000US-0216647P.
 07-JUL-2000; 2000US-0216880P.
 11-JUL-2000; 2000US-0217487P.
 14-JUL-2000; 2000US-0217496P.
 26-JUL-2000; 2000US-0218230P.
 26-JUL-2000; 2000US-0220963P.
 26-JUL-2000; 2000US-0220964P.
 14-AUG-2000; 2000US-0224518P.
 14-AUG-2000; 2000US-0224519P.
 14-AUG-2000; 2000US-0225213P.
 14-AUG-2000; 2000US-0225214P.
 14-AUG-2000; 2000US-0225266P.
 14-AUG-2000; 2000US-0225267P.
 14-AUG-2000; 2000US-0225268P.
 14-AUG-2000; 2000US-0225270P.
 14-AUG-2000; 2000US-0225447P.
 14-AUG-2000; 2000US-0225757P.
 14-AUG-2000; 2000US-0225758P.
 14-AUG-2000; 2000US-0225759P.
 18-AUG-2000; 2000US-0226279P.
 22-AUG-2000; 2000US-0226681P.
 22-AUG-2000; 2000US-0226682P.
 23-AUG-2000; 2000US-0227182P.
 30-AUG-2000; 2000US-0227009P.
 01-SEP-2000; 2000US-0228924P.
 01-SEP-2000; 2000US-0229287P.
 01-SEP-2000; 2000US-0229343P.
 01-SEP-2000; 2000US-0229344P.
 05-SEP-2000; 2000US-0229345P.
 05-SEP-2000; 2000US-0229509P.
 05-SEP-2000; 2000US-0229513P.
 06-SEP-2000; 2000US-0230437P.
 06-SEP-2000; 2000US-0230438P.
 08-SEP-2000; 2000US-0231242P.
 08-SEP-2000; 2000US-0231243P.
 08-SEP-2000; 2000US-0231244P.
 08-SEP-2000; 2000US-0231413P.
 08-SEP-2000; 2000US-0231414P.
 08-SEP-2000; 2000US-0232080P.
 08-SEP-2000; 2000US-0232081P.
 12-SEP-2000; 2000US-0231968P.
 14-SEP-2000; 2000US-0232397P.
 14-SEP-2000; 2000US-0232398P.
 14-SEP-2000; 2000US-0232399P.
 14-SEP-2000; 2000US-0232400P.
 14-SEP-2000; 2000US-0232401P.
 14-SEP-2000; 2000US-0233063P.
 14-SEP-2000; 2000US-0233064P.
 21-SEP-2000; 2000US-0233065P.
 21-SEP-2000; 2000US-0234223P.
 25-SEP-2000; 2000US-0234274P.
 25-SEP-2000; 2000US-0234997P.
 25-SEP-2000; 2000US-0234998P.
 26-SEP-2000; 2000US-0235484P.
 27-SEP-2000; 2000US-0235834P.
 27-SEP-2000; 2000US-0235835P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 02-OCT-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 02-OCT-2000; 2000US-0237040P.

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 503 AA;

Query Match 74.4%; Score 32; DB 6; Length 503;
Best Local Similarity 62.5%; Pred. No. 4.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8

Db 343 PSFINEQN 350

RESULT 45
AAM40339
ID AAM40339 standard; protein; 538 AA.

AC AAM40339;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3484.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI59495.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

XX Example 6; SEQ ID NO 3484; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

XX Sequence 538 AA;

Query Match 74.4%; Score 32; DB 4; Length 538;

Best Local Similarity 71.4%; Pred. No. 4.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7

Db 449 PDVWNIQ 455

RESULT 46

AAM40338

ID AAM40338 standard; protein; 578 AA.

AC AAM40338;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 3483.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

AC AAO19623;
 XX
 DT 13-FEB-2003 (first entry)
 XX
 DE Human nox 4 enzyme.
 XX
 KW Human; Nox; enzyme; reactive oxygen intermediate; cell division;
 KW cytotoxic; mutagenic; cancer; cardiovascular disease; NADPH oxidase;
 KW atherosclerosis; hypertension; prostatic hypertrophy; restenosis;
 KW psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200281703-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 15-NOV-2001; 2001WO-US051495.
 XX
 PR 16-NOV-2000; 2000US-0249305P.
 PR 05-DEC-2000; 2000US-0251364P.
 PR 07-MAY-2001; 2001US-0289172P.
 PR 07-MAY-2001; 2001US-0289537P.
 XX
 PA (UYEM-) UNIV EMORY.
 XX
 PI Lambeth JD, Cheng G;
 XX
 DR WPI; 2003-058555/05.
 DR N-PSDB; AAL50286.
 XX
 PT New nicotinamide adenine dinucleotide phosphate-oxidase protein,
 PT fragments, or conservative substitutions, useful for stimulating
 PT superoxide production.
 XX
 PS Claim 1; Page 68-70; 91pp; English.
 XX
 CC The present invention provides the protein and coding sequences of human
 CC NADPH oxidases designated Nox 4 and Nox 5. These proteins are capable of
 CC stimulating superoxide production, and thus are involved in the
 CC production of reactive oxygen intermediates. The proteins are therefore
 CC useful in the treatment of cancers, psoriasis, cardiovascular disease,
 CC atherosclerosis, hypertension, restenosis and prostatic hypertrophy. The
 CC present sequence is the human nox 4 enzyme
 XX
 SQ Sequence 578 AA;
 Query Match 74.4%; Score 32; DB 6; Length 578;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PSYVNVQ 7
 Db 489 PDYVNIQ 495
 RESULT 49
 ADD48774
 ID ADD48774 standard; protein; 578 AA.
 AC
 AC ADD48774;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX Human Protein NP_058627, SEQ ID NO 14484.
 DE
 XX Human; Pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; Chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.

XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; NP_058627.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 578 AA;
 Query Match 74.4%; Score 32; DB 7; Length 578;
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PSYVNVQ 7
 Db 489 PDYVNIQ 495
 RESULT 50
 ABP74053
 ID ABP74053 standard; protein; 622 AA.
 XX
 AC ABP74053;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE Candida albicans essential protein SEQ ID NO 7890.
 XX
 KW Fungus; Yeast; tetracycline; promoter; GRACE strain; biosynthesis;
 KW signal transduction; DNA replication; cell division; growth;

```

Query Match      74.4%;   Score 32;   DB 5;   Length 622;
Best Local Similarity 62.5%;   Pred. No. 5.4e+02;
Matches 5;   Conservative 2;   Mismatches 1;   Indels 0;   Gaps 0;

QY      1 PSRVNVQN 8 .
        | : | | |
Db      600 PNYLVN 607

```

Search completed: May 24, 2004, 17:31:03
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:26:17 ; Search time 21 Seconds
(without alignments)
36.644 Million cell updates/sec

Title: US-09-977-349-5
Perfect score: 43
Sequence: 1 PSYVNVQN 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : PIR 78:**

1: PIR1:**
2: PIR2:**
3: PIR3:**
4: PIR4:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	100.0	473	2 S25776	transforming prote
2	42	97.7	469	2 A55484	p52(Shc) protein -
3	36	83.7	1009	2 S20538	chitin synthase (E
4	35	81.4	858	2 JC2308	chitin synthase (E
5	35	81.4	1131	2 A23944	chitin synthase (E
6	34	79.1	187	2 S48963	hypothetical prote
7	34	79.1	392	2 T40624	probable class v p
8	33	76.7	527	2 F69378	conserved hypothet
9	33	76.7	548	2 T15318	hypothetical prote
10	33	76.7	577	2 E90044	hypothetical prote
11	33	76.7	588	2 D70394	hypothetical prote
12	33	76.7	776	2 E90591	mannosyltransferas
13	33	76.7	896	1 GNLJGH	lipoprotein limpor
14	33	76.7	896	1 GNLJGN	pol polyprotein -
15	33	76.7	910	2 B96770	hypothetical prote
16	33	76.7	990	2 E30433	conserved hypothet
17	33	76.7	1007	2 A30093	beta-galactosidase
18	32	74.4	215	2 T15811	hypothetical prote
19	32	74.4	319	2 E84980	hypothetical prote
20	32	74.4	503	2 S73843	hypothetical prote
21	32	74.4	792	2 S03232	general amino acid
22	32	74.4	858	2 JC2309	hypothetical prote
23	32	74.4	983	2 A59054	chitin synthase (E
24	32	74.4	1146	2 S64402	probable transcrip
25	31	72.1	210	2 F72645	hypothetical prote
26	31	72.1	239	2 D81218	conserved hypothet
27	31	72.1	239	2 F81795	probable periplasm
28	31	72.1	243	2 F90197	hypothetical prote
29	31	72.1	263	2 G82259	Probable rare lipo

30	31	72.1	273	2 B81211	2,3,4,5-tetrahydro
31	31	72.1	273	2 F81787	2,3,4,5-tetrahydro
32	31	72.1	274	1 XNEGSD	2,3,4,5-tetrahydro
33	31	72.1	274	1 S18857	2,3,4,5-tetrahydro
34	31	72.1	274	2 H90649	hypothetical prote
35	31	72.1	274	2 H85500	hypothetical prote
36	31	72.1	274	2 AH0127	hypothetical prote
37	31	72.1	274	2 AG0528	2,3,4,5-tetrahydro
38	31	72.1	282	2 B64063	hypothetical prote
39	31	72.1	303	1 H64133	2,3,4,5-tetrahydro
40	31	72.1	342	2 D83146	hypothetical prote
41	31	72.1	348	2 JN0916	fibbrilin precurs
42	31	72.1	348	2 JN0318	fibbrilin precurs
43	31	72.1	360	2 AF0317	rare lipoprotein A
44	31	72.1	396	1 D69053	probable iron-sulf
45	31	72.1	405	2 T01218	hypothetical prote
46	31	72.1	409	2 T11901	NADH2 dehydrogenas
47	31	72.1	429	2 S27924	gene LF2 protein -
48	31	72.1	453	2 T21528	hypothetical prote
49	31	72.1	473	2 D81322	outer membrane pro
50	31	72.1	596	2 AD1876	hypothetical prote
51	31	72.1	657	2 A11882	nitrate transport
52	31	72.1	670	2 T02019	callus-associated
53	31	72.1	760	2 S55520	chitin synthase (E
54	31	72.1	778	2 T31037	hypothetical prote
55	31	72.1	894	2 I38047	triglyceride trans
56	31	72.1	895	2 A55413	triglyceride trans
57	31	72.1	910	2 JC4609	chitin synthase (E
58	31	72.1	1209	2 T52523	hypothetical prote
59	31	72.1	1367	2 T18466	hypothetical prote
60	31	72.1	1374	1 GNNYEC	genome polyprotein
61	30	69.8	123	2 AE1730	hypothetical prote
62	30	69.8	139	2 AF0221	conserved hypothet
63	30	69.8	140	2 D90437	conserved hypothet
64	30	69.8	155	2 T22815	hypothetical prote
65	30	69.8	228	1 KTBEC	thymidine kinase (
66	30	69.8	263	2 S9380	probable membrane
67	30	69.8	274	2 H84956	2,3,4,5-tetrahydro
68	30	69.8	284	2 D95109	conserved hypothet
69	30	69.8	297	2 G97977	conserved hypothet
70	30	69.8	332	2 T20107	hypothetical prote
71	30	69.8	383	2 G96629	hypothetical prote
72	30	69.8	400	2 T28414	ORF MSV253 leucine
73	30	69.8	401	2 A75388	NADH dehydrogenase
74	30	69.8	404	2 G95365	NADH2 dehydrogenas
75	30	69.8	405	2 T42927	hypothetical prote
76	30	69.8	414	2 T50010	hypothetical prote
77	30	69.8	434	2 T26275	hypothetical prote
78	30	69.8	440	2 E84942	hypothetical prote
79	30	69.8	448	2 C64437	cell cycle protein
80	30	69.8	471	2 S57551	probable urease -
81	30	69.8	565	1 HMIVDE	hypothetical prote
82	30	69.8	573	2 C86806	hemagglutinin prec
83	30	69.8	623	2 T28051	hypothetical prote
84	30	69.8	740	2 S43566	hypothetical prote
85	30	69.8	754	2 S53373	R01H10.8 protein (
86	30	69.8	754	2 S41391	gelsolin - America
87	30	69.8	754	2 A97681	gelsolin - America
88	30	69.8	853	2 AH2905	hypothetical prote
89	30	69.8	887	2 A46764	conserved hypothet
90	30	69.8	905	1 P3XRA4	microsomal triglyc
91	30	69.8	980	2 E72035	inner capsid prote
92	30	69.8	980	2 E86589	conserved hypothet
93	30	69.8	1068	2 H96769	CM590 hypothetical
94	30	69.8	1108	2 JC4037	hypothetical prote
95	30	69.8	1203	2 T04294	alpha-mannosidase
96	30	69.8	1633	2 T01879	hypothetical prote
97	30	69.8	1802	2 G71616	hypothetical prote
98	30	69.8	3016	2 S77300	hypothetical prote
99	29	67.4	93	2 S31177	outer membrane pro
100	29	67.4	118	2 S01253	hypothetical prote
101	29	67.4	136	2 B96504	protein P9C16.21 I
102	29	67.4	162	2 D90502	hypothetical prote

103 29 67.4 220 2 A70147
 104 29 67.4 241 1 PYNVNV
 105 29 67.4 245 1 PYNVSM
 106 29 67.4 245 1 JQ1607
 107 29 67.4 245 1 PYNVPM
 108 29 67.4 245 2 S18945
 109 29 67.4 245 2 T10272
 110 29 67.4 245 2 T41754
 111 29 67.4 246 1 B40618
 112 29 67.4 250 2 AB1064
 113 29 67.4 262 2 F81826
 114 29 67.4 270 2 F83784
 115 29 67.4 282 2 S23040
 116 29 67.4 285 2 G68875
 117 29 67.4 303 2 T43411
 118 29 67.4 304 2 T23184
 119 29 67.4 329 2 T32272
 120 29 67.4 343 2 S55087
 121 29 67.4 347 2 S15171
 122 29 67.4 355 2 S15174
 123 29 67.4 367 2 E96796
 124 29 67.4 384 2 F64438
 125 29 67.4 416 2 F96690
 126 29 67.4 448 1 Q4ADA7
 127 29 67.4 449 1 Q4ADA5
 128 29 67.4 449 1 Q4ADA2
 129 29 67.4 473 2 S48563
 130 29 67.4 513 2 S50915
 131 29 67.4 529 2 D72253
 132 29 67.4 558 1 H64046
 133 29 67.4 578 2 C97167
 134 29 67.4 624 2 T49366
 135 29 67.4 637 2 T13432
 136 29 67.4 645 2 T16078
 137 29 67.4 657 2 H86344
 138 29 67.4 662 2 S42826
 139 29 67.4 690 2 S64646
 140 29 67.4 718 2 T40823
 141 29 67.4 783 2 AE1961
 142 29 67.4 816 2 H85028
 143 29 67.4 880 2 T38083
 144 29 67.4 888 2 A38539
 145 29 67.4 926 2 T39664
 146 29 67.4 950 2 C38777
 147 29 67.4 971 2 D70128
 148 29 67.4 1026 2 A49750
 149 29 67.4 1039 2 T22982
 150 29 67.4 1085 2 C96797

ALIGNMENTS

RESULT 1
 S25776
 C:Species: transforming protein (SHC) - human
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
 C:Accession: S25776
 R:Bellicci, G.; Lanfranccone, L.; Grignani, F.; McGlade, J.; Cavallo, F.; Fornì, G.; Nicol
 Cell 70, 93-104, 1992
 A>Title: A novel transforming protein (SHC) with an SH2 domain is implicated in mitogeni
 A:Reference number: S25776; MUID:9233554; PMID:1623525
 A:Accession: S25776
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-473 <PEL>
 A:Cross-references: EMBL:X68148; NID:G36453; PIDN:CAA48251.1; PID:G36454
 C:Superfamily: SH2 homology
 F:378-467/Domain: SH2 homology <SH2B>

Query Match 100.0%; Score 43; DB 2; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSYVNVQN 8
 DB 315 PSYVNVQN 322
 RESULT 2
 A55484
 p52 (Shc) protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
 C:Accession: A55484
 R:Blakie, P.; Immanuel, D.; Wu, J.; Li, N.; Yajnik, V.; Margolis, B.
 J. Biol. Chem. 269, 32031-32034, 1994
 A>Title: A region in Shc distinct from the SH2 domain can bind tyrosine-phosphorylated s
 A:Reference number: A55484; MUID:95096035; PMID:7798194
 A:Accession: A55484
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-469 <BLA>
 A:Cross-references: GB:U15784; NID:G558998; PIDN:AAC52146.1; PID:G558999
 C:Superfamily: SH2 homology
 F:374-463/Domain: SH2 homology <SH2B>

Query Match 97.7%; Score 42; DB 2; Length 469;
 Best Local Similarity 87.5%; Pred. No. 0.75;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 DB 311 PSYVNVQN 318

RESULT 3
 S20538
 chitin synthase (EC 2.4.1.16) CHS2 - yeast (Candida albicans)
 C:Species: Candida albicans
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
 C:Accession: S20538; A38192
 R:Chen-Wu, J.L.; Zwicker, J.; Bowen, A.R.; Robbins, P.W.
 Mol. Microbiol. 6, 497-502, 1992
 A>Title: Expression of chitin synthase genes during yeast and hyphal growth phases of Ca
 A:Reference number: S20538; MUID:92219994; PMID:1560778
 A:Accession: S20538
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1009 <CHE>
 A:Cross-references: GB:M82937; NID:G170844; PIDN:AAB59308.1; PID:G170845
 R:Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszló, P.J.; Robbins, P.W.
 Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
 A>Title: Classification of fungal chitin synthases.
 A:Reference number: A38192; MUID:92115692; PMID:1731323
 A:Accession: A38192
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 323-463, 'FH', '466-510', 'AS' <BOW>
 A>Note: sequence extracted from NCBI backbone (NCBI:75822); this publication is not cit
 C:Superfamily: chitin synthase chsa
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 83.7%; Score 36; DB 2; Length 1009;
 Best Local Similarity 87.5%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 DB 840 PSYVNVQN 847

RESULT 4
 JC2308
 chitin synthase (EC 2.4.1.16) 1 - Rhizopus oligosporus

C;Species: Rhizopus oligosporus
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
 C;Accession: JC2308
 R;Motoyama, T.; Sudoh, M.; Horiuchi, H.; Ohta, A.; Takagi, M.
 Biosci. Biotechnol. Biochem. 58, 1685-1693, 1994
 A;Title: Isolation and characterization of two chitin synthase genes of Rhizopus oligosporus
 A;Reference number: JC2308; MUID:95036875; PMID:7765484
 A;Accession: JC2308
 A;Molecule type: DNA
 A;Residues: 1-858 <MOT>
 A;Cross-references: DDBJ:D10159; NID:g218030; PIDN:BA01023.1; PID:g218031
 C;Genetics:
 A;Gene: chs1
 A;Introns: 118/1; 157/3; 240/3
 C;Function:
 A;Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosamine
 C;Superfamily: chitin synthase chsA
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 81.4%; Score 35; DB 2; Length 858;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 Db 697 PSYINVLN 704
 |||:|||||

RESULT 5
 A23944
 chitin synthase (EC 2.4.1.16) CHS1 - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein N1404; protein YNL192W
 C;Species: Saccharomyces cerevisiae
 C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 20-Jun-2000
 C;Accession: A23944; S63147
 R;Bullawa, C.E.; Slater, M.; Cabib, E.; Au-Young, J.; Sburlati, A.; Adair Jr., W.L.; Robb
 Cell 46, 213-225, 1986
 A;Title: The S. cerevisiae structural gene for chitin synthase is not required for chitin
 A;Reference number: A23944; MUID:86245067; PMID:2941152
 A;Accession: A23944
 A;Molecule type: DNA
 A;Residues: 1-1131 <BUL>
 A;Cross-references: EMBL:M14045; NID:g171214; PIDN:AAA34491.1; PID:g171215
 R;Obermaier, B.; Firavanti, E.; Rinke, M.; Domdey, H.
 submitted to the Protein Sequence Database, April 1996
 A;Reference number: S63122
 A;Accession: S63147
 A;Molecule type: DNA
 A;Residues: 1-1131 <OBE>
 A;Cross-references: EMBL:Z71468; NID:g1302184; PIDN:CAA96086.1; PID:g1302185; MIPS:YNL19
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: SGD:CHS1
 A;Cross-references: SGD:S0005136; MIPS:YNL192W
 A;Map position: 14L
 C;Function:
 A;Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosamine
 C;Superfamily: chitin synthase chsA
 C;Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein
 F;789-805/Domain: transmembrane #status predicted <TM1>
 F;809-825/Domain: transmembrane #status predicted <TM2>
 F;836-852/Domain: transmembrane #status predicted <TM3>
 F;866-882/Domain: transmembrane #status predicted <TM4>
 F;914-930/Domain: transmembrane #status predicted <TM5>
 F;1042-1058/Domain: transmembrane #status predicted <TM6>
 F;1100-1116/Domain: transmembrane #status predicted <TM7>

Query Match 81.4%; Score 35; DB 2; Length 1131;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||:|||||

Db 950 PSYINVLN 957

RESULT 6
 S48963
 hypothetical protein YHR121W - yeast (Saccharomyces cerevisiae)
 C;Species: Saccharomyces cerevisiae
 C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 19-Apr-2002
 C;Accession: S48963
 R;Jatrelle, P.
 submitted to the EMBL Data Library, May 1994
 A;Description: The sequence of S. cerevisiae cosmid 8263.
 A;Reference number: S46676
 A;Accession: S48963
 A;Molecule type: DNA
 A;Residues: 1-187 <LAT>
 A;Cross-references: EMBL:U00059; NID:g529116; PIDN:AAB68870.1; PID:g529138; MIPS:YHR121W
 C;Genetics:
 A;Cross-references: SGD:S0001163
 A;Map position: 8R
 C;Superfamily: Saccharomyces cerevisiae hypothetical protein YHR121W

Query Match 79.1%; Score 34; DB 2; Length 187;
 Best Local Similarity 85.7%; Pred. No. 12;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 Db 81 PSYVNV 87
 |||:|||||

RESULT 7
 T40624
 probable class v pyridoxal phosphate dependent aminotransferase - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T40624
 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
 submitted to the EMBL Data Library, December 1998
 A;Reference number: Z21941
 A;Accession: T40624
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-392 <LYN>
 A;Cross-references: EMBL:AL034563; PIDN:CAA22532.1; GSPDB:GN00067; SPDB:SPBC660.12C
 A;Experimental source: strain 972h-; cosmid c660
 C;Genetics:
 A;Gene: SPDB:SPBC660.12C
 A;Map position: 2

Query Match 79.1%; Score 34; DB 2; Length 392;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 Db 19 PDYVNVNN 26
 |||:|||||

RESULT 8
 F69378
 conserved hypothetical protein AF1030 - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
 C;Accession: F69378
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.A.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
 A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: F69378
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-527 <KLE>
A;Cross-references: GB:AE001032; GB:AE000782; NID:G2689355; PIDN:AA890213.1; PID:G264956
C;Superfamily: conserved hypothetical protein MJ1429

Query Match 76.7%; Score 33; DB 2; Length 527;
Best Local Similarity 75.0%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
DB 329 PSYLVNVN 336

RESULT 9
T15318
hypothetical protein B0304.3 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Feb-2001
C;Accession: T15318
R;Geisel, C.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of *C. elegans* cosmid B0304.
A;Reference number: Z18329
A;Accession: T15318
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-548 <GRL>
A;Cross-references: EMBL:U39472; NID:G1041877; PID:G1041879; PIDN:AAA80134.1; CBSP:B0304
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CBSP:B0304.3
A;Intons: 29/3; 73/1; 102/3; 134/1; 175/3; 321/2
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;332-493/Domain: cytochrome P450 homology <P45>
F;471/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 76.7%; Score 33; DB 2; Length 548;
Best Local Similarity 62.5%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
DB 68 PSFINVNN 75

RESULT 10
E90044
hypothetical protein SA2216 [imported] - *Staphylococcus aureus* (strain N315)
C;Species: *Staphylococcus aureus*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E90044
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E90044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-577 <KUR>
A;Cross-references: GB:BA000018; PID:G13702377; PIDN:BA843518.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA2216
C;Superfamily: beta-1,2-glucan export protein chvA; ATP-binding cassette homology

Query Match 76.7%; Score 33; DB 2; Length 577;
Best Local Similarity 75.0%; Pred. No. 71;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
DB 188 PAYHNVQN 195

RESULT 11
D70394
mannosyltransferase A - *Aquifex aeolicus*
C;Species: *Aquifex aeolicus*
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: D70394
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70394
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-588 <AQF>
A;Cross-references: GB:AE000723; NID:G2983569; PIDN:AAC07142.1; PID:G2983570; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: mufa

Query Match 76.7%; Score 33; DB 2; Length 588;
Best Local Similarity 62.5%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
DB 243 PSYITVEN 250

RESULT 12
E90591
lipoprotein [imported] - *Mycoplasma pulmonis* (strain UAB CTIP)
C;Species: *Mycoplasma pulmonis*
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C;Accession: E90591
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulm*
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: E90591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-776 <KUR>
A;Cross-references: GB:AL445566; PID:G14090052; PIDN:CAC13810.1; GSPDB:GN00153
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPV 6370
A;Genetic code: SGC3

Query Match 76.7%; Score 33; DB 2; Length 776;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
DB 394 PYFNVQN 401

RESULT 13
GNLJGH
pol polyprotein - human T-cell lymphotropic virus type 1
N;Contains: nuclease (EC 3.1.-.-); proteinase (EC 3.4.-.-); RNA-directed DNA polymerase
C;Species: human T-cell lymphotropic virus type 1, HTLV-1
A;Note: host *Homo sapiens* (man)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 16-Jul-1999
C;Accession: A03961; S02391

R;Seiki, M.; Hattori, S.; Hirayama, Y.; Yoshida, M.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3618-3622, 1983
 A;Title: Human adult T-cell leukemia virus: complete nucleotide sequence of the provirus
 A;Reference number: A93954; MUID:83221647; PMID:6304725
 A;Accession: A03961
 A;Molecule type: DNA
 A;Residues: 1-896 <SBI>
 A;Cross-references: GB:J02029; GB:M33896; NID:9425135; PIDN:AAA96673.1; PID:G331148
 A;Experimental sources: strain ATK
 A;Note: the authors translated the codon CGT for residue 15 as Leu and GAG for residue 3
 R;Bangham, C.R.M.; Danke, S.; Phillips, R.E.; Cruickshank, J.K.; Bell, J.I.
 EMBO J. 7, 4179-4184, 1988
 A;Title: Enzymatic amplification of exogenous and endogenous retroviral sequences from D
 A;Reference number: S02391; MUID:89210803; PMID:2468487
 A;Accession: S02391
 A;Molecule type: DNA
 A;Residues: 69-185 <BAN>
 A;Cross-references: EMBL:X14144; NID:G61567; PIDN:CAA32360.1; PID:G930258
 A;Note: 116-Lys was also found
 C;Comment: This protein is synthesized as a gag-pol polyprotein.
 C;Comment: The pol polyprotein contains reverse transcriptase and possibly a nuclease or
 C;Genetics:
 A;Gene: pol
 C;Superfamily: pol polyprotein
 C;Keywords: hydrolase; nucleotidyltransferase; polyprotein; reverse transcriptase

Query Match 76.7%; Score 33; DB 1; Length 896;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||:|
 Db 715 PSYINTDN 722

RESULT 14
 GNLJCN
 pol polyprotein - human T-cell lymphotropic virus type 1 (isolate Caribbean)
 N;Contains: endonuclease (EC 3.1.-.-); RNA-directed DNA polymerase (EC 2.7.7.49)
 C;Species: human T-cell lymphotropic virus type 1, HTLV-1
 A;Note: host Homo sapiens (man)
 C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jun-2000
 C;Accession: C28136
 R;Malik, K.T.A.; Even, J.; Karpas, A.
 J. Gen. Virol. 69, 1695-1710, 1988
 A;Title: Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia
 bers of the ATL/HTLV-I subgroup.
 A;Reference number: A92797; MUID:88274338; PMID:2899128
 A;Accession: C28136
 A;Molecule type: DNA
 A;Residues: 1-896 <MAL>
 A;Cross-references: GB:D13784; GB:D00294; NID:G221866; PIDN:BAA02931.1; PID:G221869
 C;Comment: The pol polyprotein contains reverse transcriptase and possibly a nuclease or
 C;Genetics:
 A;Gene: pol
 C;Superfamily: pol polyprotein
 C;Keywords: hydrolase; nucleotidyltransferase; polyprotein; reverse transcriptase

Query Match 76.7%; Score 33; DB 1; Length 896;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||:|
 Db 715 PSYINTDN 722

RESULT 15
 B96770
 hypothetical protein F1017.14 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: B96770

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: B96770
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-910 <STO>
 A;Cross-references: GB:AE005173; NID:G6939235; PIDN:AAF31737.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F1017.14
 A;Map position: 1

Query Match 76.7%; Score 33; DB 2; Length 910;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||:|
 Db 590 PEFINIQN 597

RESULT 16
 E90433
 conserved hypothetical protein [imported] - Sulfolobus solfataricus
 C;Species: Sulfolobus solfataricus
 C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C;Accession: E90433
 R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A;Description: Sulfolobus solfataricus complete genome.
 A;Reference number: A99139
 A;Accession: E90433
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-990 <KUR>
 A;Cross-references: GB:AB006641; NID:G13815906; PIDN:AAK42724.1; GSPDB:GN00155
 C;Genetics:
 A;Gene: SSO2602

Query Match 76.7%; Score 33; DB 2; Length 990;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||:|
 Db 866 PKYVNVISN 873

RESULT 17
 A30093
 beta-galactosidase (EC 3.2.1.23) - Lactobacillus delbrueckii subsp. bulgaricus
 N;Alternate names: lactase
 C;Species: Lactobacillus delbrueckii subsp. bulgaricus
 C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 15-Oct-1999
 C;Accession: A30093
 R;Schmidt, B.F.; Adams, R.M.; Requadt, C.; Power, S.; Mainzer, S.E.
 J. Bacteriol. 171, 625-635, 1989
 A;Title: Expression and nucleotide sequence of the Lactobacillus bulgaricus beta-galactos
 A;Reference number: A30093; MUID:89123132; PMID:2492511
 A;Accession: A30093
 A;Molecule type: DNA
 A;Residues: 1-1007 <SCH>
 A;Cross-references: GB:M29530; NID:G149546; PIDN:AAA25240.1; PID:G149547

C:Superfamily: beta-galactosidase
C:Keywords: glycosidase; hydrolase

Query Match 76.7%; Score 33; DB 2; Length 1007;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|:
Db 101 PQYVNVQ 107

RESULT 18

Ti5811

hypothetical protein C46H3.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15811

R:Avarello, T.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of *C. elegans* cosmid C46H3.

A:Reference number: Z18409

A:Accession: T15811

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-215 <FAV>

A:Cross-references: EMBL:U41271; NID:gl086855; PID:gl086856; PIDN:AAA82443.1; CESP:C46H3

C:Genetics:

A:Gene: CESP:C46H3.1

A:Introns: 19/1; 71/1; 99/1; 126/1

Query Match

74.4%; Score 32; DB 2; Length 215;

Best Local Similarity 75.0%; Pred. No. 37;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8

|:|:|:|:
Db 111 PNLVNVQ 118

RESULT 19

B84980

hypothetical protein [imported] - *Buchnera* sp. (strain APS)

C:Species: *Buchnera* sp.

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: B84980

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A>Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A

A:Reference number: A84930; MUID:20445173; PMID:10993077

A:Accession: B84980

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-319 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: ygfZ; BU435

Query Match

74.4%; Score 32; DB 2; Length 319;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8

|:|:|:|:
Db 2 PSFISIQ 9

RESULT 20

S73843

General amino acid permease GAP1 homolog F10_orf503 - *Mycoplasma pneumoniae* (strain ATCC

N:Alternace names: hypothetical protein F10_orf503

C:Species: *Mycoplasma pneumoniae*

A:Variety: ATCC 29342

C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73843

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A>Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*

A:Reference number: S73327; MUID:97105885; PMID:8948633

A:Accession: S73843

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-503 <HIM>

A:Cross-references: EMBL:AE000051; GB:U00089; NID:G1674211; PIDN:AAB96165.1; PID:G167421

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:

A:Gene: gap1

A:Genetic code: SGC3

Query Match

74.4%; Score 32; DB 2; Length 503;

Best Local Similarity 62.5%; Pred. No. 98;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8

|:|:|:|:
Db 343 PSFINEQ 350

RESULT 21

S03232

hypothetical protein C-792 - *Sulfolobus* particle SSV1

C:Species: *Sulfolobus* particle SSV1

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999

C:Accession: S03232

R:Palm, P.; Grampp, B.; Zillig, W.; Yeats, S.; McWilliams, P.; Reiter, W.D.

submitted to the EMBL Data Library, March 1988

A:Reference number: S03211

A:Accession: S03232

A:Molecule type: DNA

A:Residues: 1-792 <PAL>

A:Cross-references: EMBL:X07234; NID:G46703; PIDN:CAA30200.1; PID:G46725

Query Match

74.4%; Score 32; DB 2; Length 792;

Best Local Similarity 75.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8

|:|:|:|:
Db 65 PSYVSVN 72

RESULT 22

JC2309

chitin synthase (EC 2.4.1.16) 2 - *Rhizopus oligosporus*

C:Species: *Rhizopus oligosporus*

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000

C:Accession: JC2309

R:Motoyama, T.; Sudoh, M.; Horiuchi, H.; Ohta, A.; Takagi, M.

Bioeci. Biotechnol. Biochem. 58, 1685-1693, 1994

A>Title: Isolation and characterization of two chitin synthase genes of *Rhizopus oligosporus*

A:Reference number: JC2308; MUID:95036875; PMID:7765484

A:Accession: JC2309

A:Molecule type: DNA

A:Residues: 1-858 <NOT>

A:Cross-references: DDBJ:D10160; NID:G218032; PIDN:BAA01024.1; PID:G218033

C:Genetics:

A:Gene: chs2

A:Introns: 120/1; 193/3

C:Function:

A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosamine

A:Superfamily: chitin synthase chsA

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match

74.4%; Score 32; DB 2; Length 858;

Best Local Similarity 75.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
 Db 702 PSYTNVLN 709

RESULT 23
 A59054
 Chitin synthase (EC 2.4.1.16) CHS1 - Emericella nidulans
 N:Alternate names: chitin synthase chsc
 C:Species: Emericella nidulans, Aspergillus nidulans
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: A59054; JC2408; G36192
 R:Horiiuchi, H.
 Submitted to DDBJ, February 1999
 A:Description: Emericella nidulans gene for chitin synthase, complete cds.
 A:Reference number: A59054
 A:Accession: A59054
 A:Molecule type: DNA
 A:Residues: 1-983 <HOR>
 A:Cross-references: GB:AB023911; DDBJ:D38409; NID:G4519180; PIDN:BAA75501.1; PID:G451918
 A:Experimental source: strain FGSC A89
 R:Motoyama, T.; Kojima, N.; Horiiuchi, H.; Ohta, A.; Takagi, M.
 Biosci. Biotechnol. Biochem. 58, 2254-2257, 1994
 A:Title: Isolation of a chitin synthase gene (chsc) of Aspergillus nidulans.
 A:Reference number: JC2408; MUID:95128032; PMID:7765719
 A:Accession: JC2408
 A:Molecule type: DNA
 A:Residues: 1-18, 'A', 'A', '20-85, 'TNLIQVTSRWTPKHTWLTIPNHTSLQ', 'LHRAMTMSASIQDISILTRMGRSFRN
 868-889, 'LCYGVQVRSLSFLVCGFWLCEW' <MOR>
 A:Cross-references: DDBJ:D38409; NID:G4519180
 R:Bowen, A.R.; Chen-Wu, J.L.; Momany, M.; Young, R.; Szaniszlo, P.J.; Robbins, P.W.
 Proc. Natl. Acad. Sci. U.S.A. 89, 519-523, 1992
 A:Title: Classification of fungal chitin synthases
 A:Reference number: A38192; MUID:92115692; PMID:1731323
 A:Accession: G38192
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 258-280, 'RS', '283-446 <BOW>
 A:Cross-references: GB:M82938; NID:G168038; PIDN:AAA33302.1; PID:G168039
 A:Note: sequence extracted from NCBI backbone (NCBIP:75829)
 C:Genetics:
 A:Gene: chsc
 A:Introns: 7/1; 66/3; 611/3; 745/3
 C:Function:
 A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
 C:Superfamily: chitin synthase chsa
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 74.4%; Score 32; DB 2; Length 983;
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
 Db 754 PFYINVLN 761

RESULT 24
 S64402
 probable transcription factor ASK10 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein G5705; protein YGR097w
 C:Species: Saccharomyces cerevisiae
 C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
 C:Accession: S64402; S63846
 R:Hernandez, K.; Weber, N.; Wipfli, P.; Schmidheini, T.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64392
 A:Accession: S64402
 A:Molecule type: DNA
 A:Residues: 1-1146 <HER>
 A:Cross-references: EMBL:Z72882; NID:G1323146; PID:G1323147; MIPS:YGR097w

A:Experimental source: strain S288C
 R:Page, N.; Sheraton, J.; Brown, J.L.; Stewart, R.S.; Bussey, H.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: S63847
 A:Accession: S63847
 A:Molecule type: DNA
 A:Residues: 1-56, 'T', '58-345, 'N', '347-463, 'R', '465-466, 'A', '468-602, 'V', '604-905, 'P', '907-1146
 A:Cross-references: EMBL:U27209; NID:9829639; PID:9829640
 R:Page, N.; Sheraton, J.; Brown, J.L.; Stewart, R.C.; Bussey, H.
 Yeast 12, 267-272, 1996
 A:Title: Identification of ASK10 as a multicopy activator of Skm7p-dependent transcription
 A:Reference number: S63846; MUID:97060018; PMID:8904339
 A:Accession: S63846
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 233-345, 'N', '347-463, 'R', '465-466, 'A', '468-517 <PAW>
 A:Cross-references: EMBL:U27209
 A:Note: the sequence of residues 1-50 is shown in line "Homolog" rather than in line "As
 C:Genetics:
 A:Gene: SGD:ASK10
 A:Cross-references: SGD:S0003329; MIPS:YGR097w
 A:Map position: 7R
 C:Function:
 A:Description: activator of SKN7 protein-dependent transcription

Query Match 74.4%; Score 32; DB 2; Length 1146;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
 Db 784 PQVIHQIN 791

RESULT 25
 F72645
 hypothetical protein APE0596 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
 C:Accession: F72645
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Maeuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: F72645
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <KAW>
 A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79566.1; PID:dl043352; PID:G510
 C:Genetics:
 A:Gene: APE0596
 C:Superfamily: Aeropyrum pernix hypothetical protein APE0596

Query Match 72.1%; Score 31; DB 2; Length 210;
 Best Local Similarity 50.0%; Pred. No. 59;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
 Db 143 PAYIDVEN 150

RESULT 26
 DB1218
 conserved hypothetical protein NMB0267 [imported] - Neisseria meningitidis (strain MC58
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 28-Apr-2003
 A:Accession: DB1218
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: D81218
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <TET>
A:Cross-references: GB:AE002383; GB:AE002098; NID:g7225484; PIDN:AAF40721.1; PID:g722549
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0267
C:Superfamily: rplA lipoprotein

Query Match 72.1%; Score 31; DB 2; Length 239;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
| | | | |
Db 97 PSYVRVTN 104

RESULT 27
F81795
Probable periplasmic protein NMA2219 [imported] - *Neisseria meningitidis* (strain Z2491)
C:Species: *Neisseria meningitidis*
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 28-Apr-2003
C:Accession: F81795
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761319
A:Accession: F81795
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <PAR>
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85430.1; PID:g738083
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2219
C:Superfamily: rplA lipoprotein

Query Match 72.1%; Score 31; DB 2; Length 239;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
| | | | |
Db 97 PSYVRVTN 104

RESULT 28
F90197
Hypothetical protein SSO0519 [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90197
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: F90197
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-243 <KUR>
A:Cross-references: GB:AE006641; NID:G13813677; PIDN:AAK40837.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO0519

Query Match 72.1%; Score 31; DB 2; Length 243;
Best Local Similarity 62.5%; Pred. No. 70;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
| | | | |
Db 72 PSYVSIYN 79

RESULT 29
G82259
Probable rare lipoprotein A VC0948 [imported] - *Vibrio cholerae* (strain N16961 serogroup
C:Species: *Vibrio cholerae*
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 28-Apr-2003
C:Accession: G82259
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bas, S.; Qin, H.; Dragoi, I.; Sellers, P
L, R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <HEI>
A:Cross-references: GB:AE004177; GB:AE003852; NID:g9655398; PIDN:AAF94110.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0948
A:Map position: 1
C:Superfamily: rplA lipoprotein

Query Match 72.1%; Score 31; DB 2; Length 263;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
| | | | |
Db 114 PSYVKVTN 121

RESULT 30
B81211
2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase NMB0335 [imported] - Neis
C:Species: *Neisseria meningitidis*
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81211
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: B81211
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-273 <TET>
A:Cross-references: GB:AE002390; GB:AE002098; NID:g7225550; PIDN:AAF40778.1; PID:g722555
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0335
C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase

Query Match 72.1%; Score 31; DB 2; Length 273;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
| | | | |
Db 125 PSYVNI 130

RESULT 31

F81787
 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117) NMA215
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
 C:Accession: F81787
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: F81787
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-273 <PAR>
 A:Cross-references: GB:AL162758; GB:AL157959; NID:G7380672; PIDN:CAB85365.1; PID:G738077
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: dapD
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match 72.1%; Score 31; DB 2; Length 273;
 Best Local Similarity 83.3%; Pred. No. 79;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PSYVNV 6
 Db 125 PSYVNI 130

RESULT 32

XNEGSD
 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117) - Esch
 N:Alternate names: succinylaminopimelate aminotransferase; succinylaminopimelate tra
 C:Species: *Escherichia coli*
 C:Date: 13-Aug-1996 #sequence_revision 17-Oct-1997 #text_change 03-Jun-2002
 C:Accession: F64740; A00601; S45231; S36255; S31963
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: F64740
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-274 <BLAT>
 A:Cross-references: GB:AE000126; GB:U00096; NID:G1786358; PIDN:AAC73277.1; PID:G1786362;
 A:Experimental source: strain K-12, substrain MGL655
 R:Richaud, C.; Richaud, F.; Martin, C.; Haziza, C.; Patte, J.C.
 J. Biol. Chem. 259, 14824-14828, 1984
 A:Title: Regulation of expression and nucleotide sequence of the *Escherichia coli* dapD g
 A:Reference number: A00601; MUID:85054973; PMID:6094577
 A:Accession: A00601
 A:Molecule type: DNA
 A:Residues: 1-30,'D',32-162,'R',164-176,'M',178-189,'L',191-274 <RIC>
 A:Cross-references: GB:K02970; NID:G145711; PIDN:AAA23667.1; PID:G145712
 R:Fujita, N.

submitted to the EMBL Data Library, January 1994
 A:Reference number: S45181
 A:Accession: S45231
 A:Molecule type: DNA
 A:Residues: 1-274 <FUDJ>
 A:Cross-references: EMBL:D26562; NID:G473770; PIDN:BAA05610.1; PID:G473821
 A:Experimental source: strain K-12, substrain W3110
 R:van Heeswijk, W.C.; Rabenberg, M.; Westerhoff, H.V.; Kahn, D.
 Mol. Microbiol. 9, 443-457, 1993
 A:Title: The genes of the glutamine synthetase adenylation cascade are not regulated b
 A:Reference number: S36254; MUID:94018640; PMID:8412694
 A:Accession: S36255
 A:Molecule type: DNA
 A:Residues: 1-14 <VAN>
 A:Cross-references: EMBL:Z21842

C:Genetics:

A:Gene: dapD
 A:Map position: 4 min
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 C:Keywords: acyltransferase; coenzyme A; diaminopimelate biosynthesis; lysine biosynthes

Query Match 72.1%; Score 31; DB 1; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNV 6
 Db 125 PSYVNI 130

RESULT 33

S18857
 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117) - Actin
 N:Alternate names: tetrahydrodipicolinate N-succinyltransferase
 C:Species: *Actinobacillus pleuropneumoniae*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S18857

R:Denich, K.; O'Hanley, P.; Lalonde, G.
 submitted to the EMBL Data Library, November 1991
 A:Description: Cloning and sequence analysis of the DAPD gene from *Actinobacillus pleuro*;
 A:Reference number: S18857

A:Accession: S18857
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-274 <DEN>
 A:Cross-references: EMBL:X63201; NID:G38946; PIDN:CAA44883.1; PID:G38947
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 C:Keywords: acyltransferase; aminotransferase; coenzyme A; diaminopimelate biosynthesis;

Query Match 72.1%; Score 31; DB 1; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNV 6
 Db 124 PSYVNI 129

RESULT 34

H90649
 hypothetical protein ECs0168 [imported] - *Escherichia coli* (strain O157:H7, substrain RU
 C:Species: *Escherichia coli*
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90649

R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
 A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: H90649
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-274 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA833591.1; PID:G13359624; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs0168

C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase

Query Match 72.1%; Score 31; DB 2; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNV 6
 Db 125 PSYVNI 130

RESULT 35

H85500
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85500
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85500
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-274 <STO>
 A:Cross-references: GB:AE005174; NID:g12512888; PIDN:AG54468.1; GSPDB:GN00145; UWGP:ZOL
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: dapD
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase

Query Match 72.1%; Score 31; DB 2; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
 |||||:
 Db 125 PSYVNI 130

RESULT 36

AH0127
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 03-Jun-2002
 C:Accession: AH0127
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 ilo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AH0127
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-274 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89883.1; PID:g15979108; GSPDB:GN00175
 C:Genetics:
 A:Gene: dapD
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match 72.1%; Score 31; DB 2; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
 |||||:
 Db 125 PSYVNI 130

RESULT 37

AG0528
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AG0528
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Mouton, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG0528
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-274 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01367.1; PID:g16501493; GSPDB:GN00176
 C:Genetics:
 A:Gene: STV0236
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase

Query Match 72.1%; Score 31; DB 2; Length 274;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
 |||||:
 Db 125 PSYVNI 130

RESULT 38

E64063
 C:Species: *Haemophilus influenzae* (strain Rd KW20)
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: E64063
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: E64063
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-282 <TIGR>
 A:Cross-references: GB:U32720; GB:L42023; NID:g1573322; PIDN:AAC22019.1; PID:g1573328; T
 C:Superfamily: conserved hypothetical protein HI0360

Query Match 72.1%; Score 31; DB 2; Length 282;
 Best Local Similarity 75.0%; Pred. No. 82;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||:
 Db 114 PTAIVNVQN 121

RESULT 39

H64133
 C:Species: *Haemophilus influenzae*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: H64133
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: H64133
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-303 <TIGR>
 A:Cross-references: GB:U32836; GB:L42023; NID:g1574473; PIDN:AAC23279.1; PID:g1574480; T
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 C:Keywords: acyltransferase; coenzyme A; diamino-pimelate biosynthesis; lysine biosynthes

Query Match 72.1%; Score 31; DB 1; Length 303;

Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 153 PSYVNI 158

RESULT 40

DB3146
Hypothetical protein PA4000 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 28-Apr-2003
C:Accession: D83146
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83146
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <STO>
A:Cross-references: GB:AE004817; GB:AE004091; NID:g9950181; PIDN:AAG07387.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4000
C:Superfamily: rplA lipoprotein

Query Match 72.1%; Score 31; DB 2; Length 342;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNV 8
Db 137 PSYVRTN 144

RESULT 41

JN0916
fimbriin precursor - Porphyromonas gingivalis
C:Species: Porphyromonas gingivalis
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C:Accession: JN0916; JN0917
R:Fujiwara, T.; Morishima, S.; Takahashi, I.; Hamada, S.
Biochem. Biophys. Res. Commun. 197, 241-247, 1993
A:Title: Molecular cloning and sequencing of the fimbriin gene of Porphyromonas gingivalis
A:Reference number: JN0913; MUID:94071950; PMID:7902712
A:Accession: JN0916
A:Residues: 1-348 <FUJ>
A:Cross-references: DDBJ:DI7797; NID:g456497; PIDN:BAA04623.1; PID:g509036
A:Experimental source: strain HW24D1
C:Accession: JN0917
A:Molecule type: DNA
A:Residues: 1-347 <FUJ>
A:Cross-references: DDBJ:DI7798; NID:g456498; PIDN:BAA04624.1; PID:g456499
A:Experimental source: strain OM2314
C:Genetics:
A:Gene: fimA
A:Start codon: GTG
C:Function:
A:Description: this protein mediates cell adhesion to oral tissues and is important in C
C:Superfamily: Porphyromonas fimbriin
C:Keywords: cell adhesion
F:1-9/Domain: signal sequence #status predicted <SIG>
F:10-348/Product: fimbriin #status predicted <MAT>

Query Match 72.1%; Score 31; DB 2; Length 348;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNV 8
Db 147 PSYVKNY 154

RESULT 42

JN0918
fimbriin precursor - Porphyromonas gingivalis
C:Species: Porphyromonas gingivalis
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000
C:Accession: JN0918; B60275
R:Fujiwara, T.; Morishima, S.; Takahashi, I.; Hamada, S.
Biochem. Biophys. Res. Commun. 197, 241-247, 1993
A:Title: Molecular cloning and sequencing of the fimbriin gene of Porphyromonas gingivalis
A:Reference number: JN0913; MUID:94071950; PMID:7902712
A:Accession: JN0918
A:Molecule type: DNA
A:Residues: 1-348 <FUJ>
A:Cross-references: DDBJ:DI7799; NID:g456500; PIDN:BAA04625.1; PID:g456501
A:Experimental source: strain OM2409
R:Lee, J.Y.; Sojar, H.T.; Bedi, G.S.; Genco, R.J.
Infect. Immun. 59, 383-389, 1991
A:Title: Porphyromonas (Bacteroides) gingivalis fimbriin: size, amino-terminal sequence
A:Reference number: A60275; MUID:91099990; PMID:1987052
A:Accession: B60275
A>Status: preliminary
A:Molecule type: protein
A:Residues: 11-30 <LEE>
A:Experimental source: strain A7A1-28
C:Genetics:
A:Gene: fimA
A:Start codon: GTG
C:Function:
A:Description: this protein mediates cell adhesion to oral tissues and is important in C
C:Superfamily: Porphyromonas fimbriin
C:Keywords: cell adhesion
F:1-9/Domain: signal sequence #status predicted <SIG>
F:10-348/Product: fimbriin #status predicted <MAT>

Query Match 72.1%; Score 31; DB 2; Length 348;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNV 8
Db 147 PSYVKNY 154

RESULT 43

AF0317
rare lipoprotein A precursor [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AF0317
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I.
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0317
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92845.1; PID:gl5980589; GSPDB:GN00175
C:Genetics:
A:Gene: rlpA
C:Superfamily: rplA lipoprotein

Query Match 72.1%; Score 31; DB 2; Length 360;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;


```

QY      1 PSYVNVQN 8
Db      113 PSYVRVTN 120

RESULT 44
D69053
Probable iron-sulfur protein MTH140 [similarity] - Methanobacterium thermoautotrophicum
C;Species: Methanobacterium thermoautotrophicum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: D69053
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vacaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: D69053
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-396 <MTH>
A;Cross-references: GB:A8000803; GB:A8000666; NID:G2621179; PIDN:AAB84646.1; PID:G262118
A;Experimental source: strain Delta H
C;Genetics:
A;Start codon: TTG
C;Superfamily: conserved hypothetical iron-sulfur protein HP0138; ferredoxin 2 [4Fe-4S] h
F;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein
F;287-363/Domain: ferredoxin 2[4Fe-4S] homology <FER>
F;294,297,300,355/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F;304,345,348,351/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match      72.1%; Score 31; DB 1; Length 396;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
Db      246 PSYINV 251

RESULT 45
T01218
hypothetical protein F6N23.19 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 19-May-2000
C;Accession: T01218
R;Geisels, C.
submitted to the EMBL Data Library, April 1998
A;Description: The sequence of A. thaliana F6N23.
A;Reference number: Z14281
A;Accession: T01218
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-405 <GEI>
A;Cross-references: EMBL:AF058919; NID:G3047100; PID:G3047126; GSPDB:GN00063; ATSP:F6N23
C;Genetics:
A;Gene: ATSP:F6N23.19
A;Map position: 5
A;introns: 39/1; 75/3; 113/2; 156/2; 263/1; 369/2
C;Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo
F;64-118/Domain: myb DNA-binding repeat homology <MYB>

Query Match      72.1%; Score 31; DB 2; Length 405;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
Db      4 PSYINV 9

```

```

RESULT 46
T11901
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NQ04 - Thermus aquaticus thermophilu
N;Alternate names: NADH dehydrogenase I, subunit NQ04
C;Species: Thermus aquaticus thermophilus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C;Accession: T11901
R;Yano, T.; Chu, S.S.; Sled, V.S.; Ohnishi, T.; Yagi, T.
submitted to the EMBL Data Library, March 1996
A;Description: Cloning, sequencing, and expression studies of the proton-translocating N
A;Reference number: Z17372
A;Accession: T11901
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-409 <YAN>
A;Cross-references: EMBL:U52917; NID:G1279860; PID:G1279864; PIDN:AAA97941.1
C;Genetics:
A;Gene: NQ04
C;Superfamily: NADH dehydrogenase (ubiquinone) 49K protein
C;Keywords: electron transfer; membrane-associated complex; NAD; oxidoreductase

Query Match      72.1%; Score 31; DB 2; Length 409;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      373 PSFVNLQS 380

RESULT 47
S27924
gene LP2 protein - human herpesvirus 4
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jan-2000
C;Accession: S27924
R;Parker, B.D.; Bankier, A.; Satchwell, S.; Barrell, B.; Farrell, P.J.
submitted to the EMBL Data Library, August 1990
A;Description: Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B9
A;Reference number: S27923
A;Accession: S27924
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <PAR>
A;Cross-references: EMBL:M35547; NID:G330420; PIDN:AAA45897.1; PID:G330422
C;Superfamily: salivirine herpesvirus 1 hypothetical protein KCRF4

Query Match      72.1%; Score 31; DB 2; Length 429;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      48 PSYVRVTN 55

RESULT 48
T21528
hypothetical protein F28H7.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21528
R;Berks, M.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19435
A;Accession: T21528
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-453 <WIL>
A;Cross-references: EMBL:Z72508; PIDN:CAA96639.1; GSPDB:GN00023; CESP:F28H7.8
A;Experimental source: clone F28H7
C;Genetics:
A;Gene: CESP:F28H7.8

```

Job time : 35 secs

A;Map position: 5
A;Introns: 145/3; 177/1; 210/2; 291/1; 410/2

Query Match 72.1%; Score 31; DB 2; Length 453;
Best Local Similarity 83.3%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
|||:|
Db 203 PSYINV 208

RESULT 49
D81322
outer membrane protein Cj1170c [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: D81322
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: AB1250; MUID:20150912; PMID:10688204
A;Accession: D81322
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-473 <PAR>
A;Cross-references: GB:AL111168; GB:AL139077; PID:96968444; PIDN:CAB73424.1; PID:9696860
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1170c
C;Superfamily: Campylobacter jejuni outer membrane protein Cj1170c

Query Match 72.1%; Score 31; DB 2; Length 473;
Best Local Similarity 71.4%; Pred. NO. 1.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SYVNVQN 8
|||:|
Db 437 SYTNQN 443

RESULT 50
AD1876
hypothetical protein alr0557 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD1876
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD1876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-596 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA072515.1; PID:GJ129902; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0557

Query Match 72.1%; Score 31; DB 2; Length 596;
Best Local Similarity 83.3%; Pred. NO. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YVNVQN 8
|||:|
Db 561 YINVQN 566

Search completed: May 24, 2004, 17:32:07

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:25:22 ; Search time 12 Seconds
(without alignments)

34.713 Million cell updates/sec

Title: US-09-977-349-5

Perfect score: 43

Sequence: 1 PSYVNVQN 8

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	583	1 SHC HUMAN	P29353 homo sapien
2	42	97.7	579	1 SHC MOUSE	P98083 mus musculus
3	36	83.7	1009	1 CHS2 CANAL	P30572 candida alb
4	35	81.4	858	1 CHS1 RHIOAL	P30594 rhizopus ol
5	35	81.4	1131	1 CHS1 YEAST	P08004 saccharomyc
6	35	81.4	5938	1 MAC4 HUMAN	Q96PK2 homo sapien
7	34	79.1	187	1 YHS1 YEAST	P38828 saccharomyc
8	33	76.7	267	1 ZDHC HUMAN	Q96GR4 homo sapien
9	33	76.7	896	1 POL HTL1A	P03362 human t-cel
10	33	76.7	896	1 POL HTL1C	P14078 human t-cel
11	33	76.7	1006	1 EGAL LACDE	P20043 lactobacill
12	32	74.4	319	1 Y435 BUCAL	P57510 buchnera ap
13	32	74.4	503	1 Y226 MYCPN	P75462 mycoplasma
14	32	74.4	792	1 Y85K SSV1	P20210 sulfolobus
15	32	74.4	858	1 CHS2 RHIOAL	P30595 rhizopus ol
16	32	74.4	928	1 CHS2 EXHOLE	P30601 exophiala d
17	32	74.4	1146	1 AS10 YEAST	P48361 saccharomyc
18	31	72.1	239	1 RLPA NEIMA	Q9JEM7 neisseria m
19	31	72.1	239	1 RLPA NEIMB	Q9KIA0 neisseria m
20	31	72.1	263	1 RLPA VIBCH	Q9KIF4 vibrio chol
21	31	72.1	273	1 DAPD BORPE	Q9ZEX2 bordetella
22	31	72.1	273	1 DAPD BUCAP	O85290 buchnera ap
23	31	72.1	274	1 DAPD ACTPL	P41396 actinobacil
24	31	72.1	274	1 DAPD ECOLI	P03948 escherichia
25	31	72.1	274	1 DAPD MYCBO	P56220 mycobacteri
26	31	72.1	275	1 DAPD HAEIN	P45284 haemophilus
27	31	72.1	282	1 Y360 HAEIN	P44661 haemophilus
28	31	72.1	342	1 RLPA PSEAE	Q9XSV6 pseudomonas
29	31	72.1	348	1 FMA2 PORGI	Q51822 porphyromon
30	31	72.1	360	1 RLPA YERPE	Q8ZDG6 yersinia pe
31	31	72.1	409	1 NQO4 THETH	Q56220 thermus the
32	31	72.1	453	1 YUOF CAEAL	Q19895 caenorhabdi
33	31	72.1	760	1 CHSX_USTMA	Q99126 ustilago ma

841	1	CHS1_PHYBL	P87073 phycomyces
894	1	MTP_HUMAN	P55157 homo sapien
894	1	MTP_MOUSE	O08601 mus musculus
895	1	MTP_MESAU	P55158 mesocricetu
910	1	CHSA_AMPOU	Q12564 ampelomyces
910	1	WRN_HUMAN	Q14191 homo sapien
910	1	WRN_HUMAN	P28855 ictalurid h
910	1	KITH_HSV1	Q8D193 synechococc
910	1	SFSA_SYNEL	P57323 buchnera ap
910	1	DAPD_BUCAL	P57323 buchnera ap
910	1	FAC2_DROME	Q9ULH8 drosophila
910	1	NUOD_DEIRA	Q9RU89 deinococcus
910	1	NUD2_RHIME	P56908 rhizobium m
910	1	MSJ7_BUCAL	P57211 buchnera ap
910	1	YB00_METUA	Q58500 methanococc
910	1	UBP8_YEAST	P50102 saccharomyc
910	1	HEMA_IADBI	P04661 influenza a
910	1	HPPI_METAC	Q8TJA9 methanosarc
910	1	GELS_HOMAM	Q27319 homarus ame
910	1	CHS1_RHIRA	Q12632 rhizomucor
910	1	MTP_BOVIN	P55156 bos taurus
910	1	MAN2_DROME	Q24451 drosophila
910	1	Y15C_ECOLI	P19770 escherichia
910	1	Y378_BORBU	O50165 borrelia bu
910	1	PYHD_NPVEM	P03237 bombyx mori
910	1	PYHD_NPVAG	P81472 anticarsia
910	1	PYHD_NPVHC	P23273 hyphantria
910	1	PYHD_NPVOP	P07488 orgyia pseu
910	1	SEFEP_SALEN	P33387 salmonella
910	1	ATPG_PROMO	P29710 propionigen
910	1	MAL2_SCHPO	Q10290 schizosacch
910	1	RPOA_LOTUA	Q9BBQ4 lotus japon
910	1	SRT1_YEAST	Q03175 saccharomyc
910	1	Y681_PASMG	P57864 pasteurella
910	1	PIV2_ADE07	P03273 human adeno
910	1	CCA_METJA	Q58511 methanococc
910	1	PIV2_ADE02	P03272 human adeno
910	1	PIV2_ADE05	P03271 human adeno
910	1	PIV2_ADE01	P12539 mouse adeno
910	1	TBG_YEAST	P53378 saccharomyc
910	1	VPL_BPMU	P79678 bacterioph
910	1	RECEN_HAEIN	P44496 haemophilus
910	1	YME1_SCHMA	P46508 schistosoma
910	1	BPL1_YEAST	P48445 s biotin-p
910	1	TRK2_SCHPO	Q10065 schizosacch
910	1	CHS2_SCHPO	O74756 schizosacch
910	1	Y228_BORBU	O51246 borrelia bu
910	1	BGAL_STRTR	P23989 streptococ
910	1	SECS_ARATH	Q883U9 arabidopsis
910	1	HP55_HUMAN	Q9H723 homo sapien
910	1	CNA2_HUMAN	Q05975 rickettsia
910	1	Y785_RICPR	Q57458 bacillus th
910	1	CLEA_BACTX	P47582 mycoplasma
910	1	RPOC_MYCCE	Q99683 homo sapien
910	1	M3X5_HUMAN	O35099 mus muscul
910	1	M3X5_MOUSE	O14802 homo sapien
910	1	RPC1_HUMAN	Q23868 dictyosteli
910	1	TAGC_DICDI	Q9W089 drosophila
910	1	DIP2_DROME	O60281 homo sapien
910	1	Z392_HUMAN	Q00402 saccharomyc
910	1	NUM1_YEAST	Q9XDX3 leptospira
910	1	RL23_LEPTN	P35238 homo sapien
910	1	RHOG_HUMAN	O9SPQ8 dictyosteli
910	1	RACL_DICDI	Q9PL91 chlamydia m
910	1	TRMB_CHLMU	O84836 chlamydia t
910	1	TRMB_CHLTR	P33409 bordetella
910	1	FIMB_BORPE	P24646 spodoptera
910	1	PYHD_NPVSL	P54450 bacillus su
910	1	YOEE_BACSU	P31165 piscum sativ
910	1	RK15_PEA	Q7VRE5 candidatus
910	1	EFTS_CANBF	P33895 drosophila
910	1	NUZM_DROYA	Q44297 arthrobacte
910	1	ODH_ARTSP	Q87HJ2 vibrio para
910	1	FAH2_VIBPA	

107 28 65.1 1 ALR_LACPL
108 28 65.1 1 ALR_LACRE
109 28 65.1 1 PEPC_MAFU
110 28 65.1 1 IPYR_PLAF7
111 28 65.1 380 1 ALR1_STAAM
112 28 65.1 382 1 ALR1_STAAM
113 28 65.1 382 1 ALR1_STAAM
114 28 65.1 391 1 NUCC_MESVI
115 28 65.1 391 1 NUCC_MESVI
116 28 65.1 430 1 ARCA_COREF
117 28 65.1 437 1 ARCA_COREF
118 28 65.1 446 1 PIV2_ADRCC
119 28 65.1 446 1 PIV2_ADRCC
120 28 65.1 452 1 PIV2_ADRCC
121 28 65.1 466 1 HTOA_HABIN
122 28 65.1 473 1 ALL2_ALISA
123 28 65.1 486 1 ALL1_ALISA
124 28 65.1 493 1 MOCR_RHIME
125 28 65.1 502 1 PIV2_ADRCC
126 28 65.1 506 1 VP7_RDVA
127 28 65.1 506 1 VP7_RDVA
128 28 65.1 506 1 VP7_RDVA
129 28 65.1 574 1 YJH1_YEAST
130 28 65.1 591 1 VATA_CHLPN
131 28 65.1 609 1 RFAL_SCHPO
132 28 65.1 609 1 YKD6_CAREL
133 28 65.1 613 1 TX18_MOUSE
134 28 65.1 633 1 COS2_METAC
135 28 65.1 634 1 COS2_METAC
136 28 65.1 656 1 DCHS_RAT
137 28 65.1 662 1 DCHS_MOUSE
138 28 65.1 663 1 Y104_MENJA
139 28 65.1 778 1 ARCB_ECO57
140 28 65.1 807 1 YAK1_YEAST
141 28 65.1 814 1 PHSG_CHLTR
142 28 65.1 830 1 VGLB_HSV6G
143 28 65.1 830 1 VGLB_HSV6U
144 28 65.1 830 1 VGLB_HSV6Z
145 28 65.1 897 1 APGL_YEAST
146 28 65.1 906 1 NUOG_BUCAL
147 28 65.1 908 1 TB12_NEIMB
148 28 65.1 911 1 TB11_NEIMB
149 28 65.1 915 1 TBPL_NEIGO
150 28 65.1 917 1 GCFC_HUMAN

ALIGNMENTS

RESULT 1
SHC_HUMAN
ID SHC_HUMAN STANDARD; PRT; 583 AA.
AC P29353; O15290;
DT 01-DEC-1992 (Rel. 24, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE SHC transforming protein.
GN SHC1 OR SHC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM P46SHC AND P52SHC).
RX MEDLINE=92323554; PubMed=1623525;
RA Pelicci G., Lanfrancone L., Grignani F., McGlade J., Cavallo F.,
RA Fornì G., Nicoletti I., Grignani F., Pawson T., Pelicci P.G.;
RT "A novel transforming protein (SHC) with an SH2 domain is implicated
in mitogenic signal transduction.";
RL Cell 70:93-104(1992).
[2]
SEQUENCE FROM N.A. (ISOFORM P66SHC).
RX MEDLINE=97201514; PubMed=9049300;

RA Migliaccio E., Mele S., Salcini A.E., Pelicci G., Lai K.M.,
RA Superti-Furga G., Pawson T., Di Fiore P.P., Lanfrancone L.,
RA Pelicci P.G.;
RT "Opposite effects of the p52shc/p46shc and p66shc splicing isoforms on
RT the EGF receptor-MAP kinase-fos signalling pathway.";
RL EMBO J. 16:706-716(1997).
[3]
SEQUENCE FROM N.A. (ISOFORM P66SHC).
RA Harun R., Smith K.K., Leek J.P., Markham A.F., Norris A.,
RA Morrison J.F.J.;
RN Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
[4]
INTERACTION WITH APS.
RX MEDLINE=97377002; PubMed=9233773;
RA Yokouchi M., Suzuki R., Masuhara M., Komiya S., Inoue A.,
RA Yoshimura A.;
RT "Cloning and characterization of APS, an adaptor molecule containing
RT PH and SH2 domains that is tyrosine phosphorylated upon B-cell
RT receptor stimulation.";
RL Oncogene 15:7-15(1997).
[5]
X-RAY CRYSTALLOGRAPHY (2.7 ÅNGSTROMS) OF 482-583.
RX MEDLINE=96080377; PubMed=7473762;
RA Mikol V., Baumann G., Zurini M.G.M., Hommel U.;
RT "Crystal structure of the SH2 domain from the adaptor protein SHC: a
RT model for peptide binding based on X-ray and NMR data.";
RL J. Mol. Biol. 254:86-95(1995).
[6]
STRUCTURE BY NMR OF 127-317.
RX MEDLINE=96097066; PubMed=8524391;
RA Zhou M.M., Ravichandran K.S., Olejniczak E.F., Petros A.M.,
RA Meadows R.P., Sattler M., Harlan J.E., Wade W.S., Burakoff S.J.,
RA Fesik S.W.;
RT "Structure and ligand recognition of the phosphotyrosine binding
RT domain of Shc.";
RL Nature 378:584-592(1995).
CC -I- FUNCTION: May couple activated growth factor receptors to a
CC signaling pathway that regulates the proliferation of mammalian
CC cells. SHC might participate in the transforming activity of
CC oncogenic tyrosine kinases.
CC -I- SUBUNIT: Interacts with the N-terminal region of APS.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- ALTERNATIVE PRODUCTS:
CC Event-Alternative initiation:
CC Comment=3 isoforms, p66shc (shown here), p52shc and p46shc, are
CC produced by alternative initiation;
CC -I- TISSUE SPECIFICITY: Widely expressed.
CC -I- DOMAIN: The PID domain specifically binds to the Asn-Pro-Xaa-
CC Tyr(P) motif found in many tyrosine-phosphorylated proteins
CC including growth factor receptors.
CC -I- PTM: Phosphorylated by activated epidermal growth factor receptor.
CC -I- SIMILARITY: Contains 1 SH2 domain.
CC -I- SIMILARITY: Contains 1 PID domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X68148; CAA48251.1; -;
DR EMBL; Y03377; AAB49972.1; -;
DR EMBL; Y09847; CAA70977.1; -;
DR PDB; 1MLI; 08-NOV-96.
DR PDB; 1SHC; 15-MAY-97.
DR PDB; 1QGI; 17-JUN-99.
DR Genew; HGNC:10840; SHC1.
DR MTM; 600560; -;
DR GO; GO:0005068; F:transmembrane receptor protein tyrosine kin...; TAS.
DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
DR GO; GO:0007176; P:regulation of EGF receptor activity; TAS.

DR InterPro; IPR006019; PID domain.
 DR InterPro; IPR006020; PTB_PID.
 DR InterPro; IPR000980; SH2_PID.
 DR Pfam; PF00640; PID; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00629; SHC2DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00462; PTB; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS01179; PID; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR SH2 domain; Phosphorylation; Growth regulation; Proto-oncogene;
 KW Alternative initiation; 3D-structure.
 FT CHAIN 1 583 SHC TRANSFORMING PROTEIN, ISOFORM P66SHC.
 FT CHAIN 111 583 SHC TRANSFORMING PROTEIN, ISOFORM P52SHC.
 FT CHAIN 156 583 SHC TRANSFORMING PROTEIN, ISOFORM P46SHC.
 FT INIT_MET 111 111 FOR ISOFORM P52SHC.
 FT DOMAIN 156 156 PID.
 FT DOMAIN 156 339 PID.
 FT DOMAIN 488 579 SH2.
 FT CONFLICT 2 2 D -> N (IN REF. 3).
 FT CONFLICT 21 21 L -> M (IN REF. 3).
 FT CONFLICT 38 38 S -> P (IN REF. 3).
 FT CONFLICT 95 95 D -> V (IN REF. 3).
 FT CONFLICT 101 101 D -> D (IN REF. 3).
 FT TURN 131 132
 FT TURN 150 156
 FT HELIX 157 158
 FT TURN 160 172
 FT STRAND 181 182
 FT TURN 183 198
 FT HELIX 221 237
 FT STRAND 238 239
 FT TURN 240 245
 FT TURN 246 249
 FT STRAND 250 256
 FT STRAND 262 264
 FT TURN 269 270
 FT STRAND 274 278
 FT STRAND 286 291
 FT HELIX 297 308
 FT TURN 309 310
 FT TURN 310 310
 FT TURN 483 485
 FT TURN 487 488
 FT STRAND 489 489
 FT STRAND 495 499
 FT TURN 500 501
 FT TURN 505 506
 FT STRAND 507 512
 FT STRAND 520 526
 FT TURN 527 528
 FT TURN 527 528
 FT STRAND 529 536
 FT TURN 538 539
 FT STRAND 542 543
 FT STRAND 548 549
 FT STRAND 552 562
 FT TURN 563 563
 FT STRAND 566 568
 FT TURN 569 570
 FT STRAND 571 573
 FT STRAND 577 578
 SQ SEQUENCE 583 AA; 62852 MW; 8B07E8B06711C933 CRC64;

Query Match 100.0%; Score 43; DB 1; Length 583;
 Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
 |||||
 Db 425 PSYVNVQN 432

RESULT 2
 SHC_MOUSE
 ID SHC_MOUSE STANDARD; PRT; 579 AA.
 AC P98083;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE SHC transforming protein.
 GN SHC1 OR SHC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM P66).
 RA Blaikie P.A., Yajnik V., Margolis B.;
 RL submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM P52).
 RA Blaikie P.A., Immanuel D., Wu J., Li N., Yajnik V., Margolis B.;
 RL MEDLINE=95096035; PubMed=7798194;
 RT "A region in Shc distinct from the SH2 domain can bind tyrosine-phosphorylated growth factor receptors."
 CC J. Biol. Chem. 269:32031-32034(1994).
 CC -!- FUNCTION: May couple activated growth factor receptors to a signaling pathway that regulates the proliferation of mammalian cells. SHC might participate in the transforming activity of oncogenic tyrosine kinases.
 CC -!- SUBUNIT: Interacts with the N-terminal region of APS (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative initiation;
 CC Comment=3 isoforms, p66 (shown here), p52 and p47, are produced by alternative initiation;
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- DOMAIN: The PID domain specifically binds to the Asn-Pro-Xaa-Tyr(P) motif found in many tyrosine-phosphorylated proteins including growth factor receptors.
 CC -!- PTM: Phosphorylated by activated epidermal growth factor receptor (By similarity).
 CC -!- SIMILARITY: Contains 1 SH2 domain.
 CC -!- SIMILARITY: Contains 1 PID domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U46956; AAA91777.2; -;
 DR EMBL; U15784; AAC52146.1; -;
 DR HSPSP; P29353; ISHC.
 DR MGD; MGI:98296; Shc1.
 DR GO; GO:0005515; P:protein binding; IPI.
 DR InterPro; IPR006019; PID domain.
 DR InterPro; IPR006020; PTB_PID.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00640; PID; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00629; SHC2DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00462; PTB; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS01179; PID; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR SH2 domain; Phosphorylation; Growth regulation; Proto-oncogene;
 KW Alternative initiation.
 FT CHAIN 1 579 SHC TRANSFORMING PROTEIN, ISOFORM P66.
 FT CHAIN 111 579 SHC TRANSFORMING PROTEIN, ISOFORM P52.

FT CHAIN 156 579 SHC TRANSFORMING PROTEIN, ISOFORM P47.
 FT INIT MET 111 111 FOR ISOFORM P52.
 FT INIT MET 156 156 FOR ISOFORM P47.
 FT DOMAIN 156 339 PID.
 FT DOMAIN 484 575 SH2.
 SQ SEQUENCE 579 AA; 62607 MW; 99C22E64412B6236 CRC64;

Query Match 97.7%; Score 42; DB 1; Length 579;
 Best Local Similarity 87.5%; Pred. No. 0.54;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||:
 DB 421 PSYVNIQN 428

RESULT 3
 CHS2_CANAL STANDARD; PRT; 1009 AA.
 AC P30572;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase 2).
 DE CHS2.
 GN Candida albicans (Yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219994; PubMed=1560778;
 RA Chen-Wu J.L., Zwicker J., Bowen A.R., Robbins P.W.;

RT "Expression of chitin synthase genes during yeast and hyphal growth phases of *Candida albicans*.";
 RL Mol. Microbiol. 6:497-502(1992).
 CC -!- FUNCTION: Formation and repair of the disk-shaped septum in yeast and the cross walls of the hyphal phase.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N+1).
 CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
 CC -!- DEVELOPMENTAL STAGE: Very high levels of CHS2 in cells undergoing hyphal outgrowth.
 CC -!- SIMILARITY: Belongs to the chitin synthase family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; M82937; AAB59308.2; --
 DR InterPro; IPR004834; Chitin_synth.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF01644; Chitin_synth; 1.
 DR ProDom; PD002998; Chitin_synth; 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KW Multigene family.
 SQ SEQUENCE 1009 AA; 115595 MW; 182B660678549EF9 CRC64;

Query Match 83.7%; Score 36; DB 1; Length 1009;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||:
 DB 840 PSYVNVLN 847

RESULT 4
 CHS1_RHIOL STANDARD; PRT; 858 AA.
 ID CHS1_RHIOL
 AC P30594;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase 1).
 DE CHS1.
 GN Rhizopus oligosporus.
 OS Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 OC Rhizopus.
 OX NCBI_TaxID=4847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95036875; PubMed=7765484;
 RA Motoyama T., Sudoh M., Horiuchi H., Ohta A., Takagi M.;

RT "Isolation and characterization of two chitin synthase genes of *Rhizopus oligosporus*.";
 RL Biosci. Biotechnol. Biochem. 58:1685-1693(1994).
 CC -!- FUNCTION: Plays a major role in cell wall biogenesis.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N+1).
 CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
 CC -!- SIMILARITY: Belongs to the chitin synthase family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; D10159; BAA01023.1; --
 DR PIR; JC2308; JC2308.
 DR InterPro; IPR004834; Chitin_synth.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF01644; Chitin_synth; 1.
 DR ProDom; PD002998; Chitin_synth; 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KW Multigene family.
 SQ SEQUENCE 858 AA; 97057 MW; EE1E6197F00E70B9 CRC64;

Query Match 81.4%; Score 35; DB 1; Length 858;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 |||||:
 DB 697 PSYVNVLN 704

RESULT 5
 CHS1_YEAST STANDARD; PRT; 1131 AA.
 ID CHS1_YEAST
 AC P08004;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase 1).
 DE CHS1 OR YNL192W OR N1404.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86245067; PubMed=2941152;
 RA Bulawa C.E., Slater M., Cabib E., Au-Young J., Sburlati A.,

```

RA Adair W.L. Jr., Robbins P.W.;
RT "The S. cerevisiae structural gene for chitin synthase is not
RL required for chitin synthesis in vivo.";
RN Cell 46:213-225(1986).
RP [2]
SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL submitted (MAY-1996) to the EMBL/GenBank/DDJ databases.
CC -!- FUNCTION: Septum formation and repair, especially under certain
CC adverse conditions.
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-
CC beta-D-glucosaminyl)) (N) = UDP + ((1,4)-(N-acetyl-beta-D-
CC glucosaminyl)) (N+1).
CC -!- ENZYME REGULATION: Requires proteolytic activation.
CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -!- MISCELLANEOUS: At least 4 potential membrane-spanning regions
CC exist.
CC -!- SIMILARITY: Belongs to the chitin synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14045; AAA34491.1; -.
CC EMBL; Z71468; CAA96086.1; -.
CC PIR; A23944; A23944.
CC GerMOnline; 143198; -.
CC SGD; S0005136; CHS1.
CC GO; GO:0004100; F:chitin synthase activity; IDA.
CC InterPro; IPR004834; Chitin synth.
CC InterPro; IPR001173; GlycoTrans_2.
CC Pfam; PF01644; Chitin synth; 1.
CC ProDom; PD002998; Chitin synth; 1.
CC Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
CC Multigene family.
CC SQ SEQUENCE 1131 AA; 129918 MW; BF6D3F357812E03 CRC64;

Query Match 81.4%; Score 35; DB 1; Length 1131;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PSYVNVQN 8
Db |||:|:|
950 PSYVNVIN 957

RESULT 6
MAC4 HUMAN STANDARD; PRT; 5938 AA.
AC Q96PK2; Q9WKY1.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoform 4.
GN MACF1 OR AC7F OR ABP620 OR KIAA0465 OR KIAA1251.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE-21833812; PubMed-11845298;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "MACF1 gene structure: a hybrid of plectin and dystrophin.";
RL Mamm. Genome 12:852-861(2001).
CC -!- FUNCTION: May play a role in cross-linking cytoskeletal proteins
CC by binding intermediate filaments to the N-terminal plectin
CC repeats and microtubules to the C-terminus.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
```

```

CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q96PK2-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=External;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=External;
CC Name=3;
CC IsoId=Q9UPN3-3; Sequence=External;
CC -!- TISSUE SPECIFICITY: Expressed in heart, lung, pituitary and
CC placenta, not found in brain, kidney, liver, pancreas or skeletal
CC muscle.
CC -!- SIMILARITY: Belongs to the plakin or cytolinker family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 19 plectrin repeats.
CC -!- SIMILARITY: Contains 32 spectrin repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF317696; AAL09459.1; -.
CC EMBL; AF325341; AAL38997.1; -.
CC EMBL; AF325333; AAL38997.1; JOINED.
CC EMBL; AF325334; AAL38997.1; JOINED.
CC EMBL; AF325335; AAL38997.1; JOINED.
CC EMBL; AF325336; AAL38997.1; JOINED.
CC EMBL; AF325339; AAL38997.1; JOINED.
CC EMBL; AF325340; AAL38997.1; JOINED.
CC GO; GO:0005856; C:cytoskeleton; ISS.
CC GO; GO:0005503; F:calcium ion binding; ISS.
CC GO; GO:0008017; F:microtubule binding; ISS.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR003108; GAS2.
CC InterPro; IPR001101; Plectrin repeat.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00036; ehand; 2.
CC Pfam; PF0187; GAS2; 1.
CC Pfam; PF00681; Plectrin; 11.
CC Pfam; PF00435; spectrin; 26.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00054; EFh; 2.
CC SMART; SM00243; GAS2; 1.
CC SMART; SM00250; PLEC; 19.
CC SMART; SM00150; SPEC; 32.
CC PROSITE; PS00018; EF_HAND; 2.
CC Cytoskeleton; Calcium; Calcium-binding; Repeat; Coiled coil;
CC Alternative splicing.
CC DOMAIN 1830 1936 COILED COIL (POTENTIAL).
CC DOMAIN 2001 2192 COILED COIL (POTENTIAL).
CC DOMAIN 2282 2345 COILED COIL (POTENTIAL).
CC DOMAIN 2477 2507 COILED COIL (POTENTIAL).
CC DOMAIN 2541 2654 COILED COIL (POTENTIAL).
CC DOMAIN 2686 2814 COILED COIL (POTENTIAL).
CC DOMAIN 2887 2919 COILED COIL (POTENTIAL).
CC DOMAIN 3046 3197 COILED COIL (POTENTIAL).
CC DOMAIN 3262 3503 COILED COIL (POTENTIAL).
CC DOMAIN 3632 3666 COILED COIL (POTENTIAL).
CC DOMAIN 3746 3779 COILED COIL (POTENTIAL).
CC DOMAIN 3920 3984 COILED COIL (POTENTIAL).
CC DOMAIN 4098 4168 COILED COIL (POTENTIAL).
CC DOMAIN 4288 4308 COILED COIL (POTENTIAL).
CC DOMAIN 4354 4386 COILED COIL (POTENTIAL).
CC DOMAIN 4397 4433 COILED COIL (POTENTIAL).
CC DOMAIN 4469 4489 COILED COIL (POTENTIAL).
CC DOMAIN 4586 4720 COILED COIL (POTENTIAL).
CC DOMAIN 4845 4880 COILED COIL (POTENTIAL).
CC DOMAIN 4910 4939 COILED COIL (POTENTIAL).
```

FT DOMAIN 4970 5000 COILED COIL (POTENTIAL).
 FT DOMAIN 5409 5437 COILED COIL (POTENTIAL).
 FT DOMAIN 5546 5569 COILED COIL (POTENTIAL).
 FT CA_BIND 5598 5610 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 5634 5646 EF-HAND 2 (POTENTIAL).
 FT REPEAT 12 48 PLECTIN 1.
 FT REPEAT 53 89 PLECTIN 2.
 FT REPEAT 89 126 PLECTIN 3.
 FT REPEAT 130 164 PLECTIN 4.
 FT REPEAT 166 202 PLECTIN 5.
 FT REPEAT 203 240 PLECTIN 6.
 FT REPEAT 243 278 PLECTIN 7.
 FT REPEAT 279 316 PLECTIN 8.
 FT REPEAT 318 354 PLECTIN 9.
 FT REPEAT 725 762 PLECTIN 10.
 FT REPEAT 801 838 PLECTIN 11.
 FT REPEAT 839 876 PLECTIN 12.
 FT REPEAT 897 934 PLECTIN 13.
 FT REPEAT 935 972 PLECTIN 14.
 FT REPEAT 975 1011 PLECTIN 15.
 FT REPEAT 1012 1049 PLECTIN 16.
 FT REPEAT 1121 1158 PLECTIN 17.
 FT REPEAT 1159 1196 PLECTIN 18.
 FT REPEAT 1295 1332 PLECTIN 19.
 FT REPEAT 2052 2160 SPECTRIN 1.
 FT REPEAT 2211 2313 SPECTRIN 2.
 FT REPEAT 2320 2430 SPECTRIN 3.
 FT REPEAT 2437 2543 SPECTRIN 4.
 FT REPEAT 2550 2652 SPECTRIN 5.
 FT REPEAT 2659 2758 SPECTRIN 6.
 FT REPEAT 2765 2896 SPECTRIN 7.
 FT REPEAT 2903 3008 SPECTRIN 8.
 FT REPEAT 3015 3119 SPECTRIN 9.
 FT REPEAT 3126 3229 SPECTRIN 10.
 FT REPEAT 3236 3339 SPECTRIN 11.
 FT REPEAT 3346 3446 SPECTRIN 12.
 FT REPEAT 3453 3555 SPECTRIN 13.
 FT REPEAT 3562 3664 SPECTRIN 14.
 FT REPEAT 3671 3775 SPECTRIN 15.
 FT REPEAT 3782 3884 SPECTRIN 16.
 FT REPEAT 3891 3993 SPECTRIN 17.
 FT REPEAT 4000 4102 SPECTRIN 18.
 FT REPEAT 4109 4211 SPECTRIN 19.
 FT REPEAT 4218 4320 SPECTRIN 20.
 FT REPEAT 4327 4428 SPECTRIN 21.
 FT REPEAT 4438 4544 SPECTRIN 22.
 FT REPEAT 4551 4653 SPECTRIN 23.
 FT REPEAT 4660 4763 SPECTRIN 24.
 FT REPEAT 4770 4872 SPECTRIN 25.
 FT REPEAT 4879 4982 SPECTRIN 26.
 FT REPEAT 4989 5091 SPECTRIN 27.
 FT REPEAT 5098 5201 SPECTRIN 28.
 FT REPEAT 5208 5309 SPECTRIN 29.
 FT REPEAT 5316 5418 SPECTRIN 30.
 FT REPEAT 5425 5555 SPECTRIN 31.
 FT REPEAT 5555 5712 SPECTRIN 32.
 FT CONFLICT 1712 1712 P -> S (IN REF. 1; AAL38997).
 SQ SEQUENCE 5938 AA; 670134 MW; B8784112752DA004 CRC64;
 Query Match 81.4%; Score 35; DB 1; Length 5938;
 Best Local Similarity 85.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 SYINVQN 8
 DB 411 SYINVQN 417
 RESULT 7
 YH51 YEAST
 ID YH51 YEAST STANDARD; PRT; 187 AA.
 AC P38826;
 DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 21.3 kDa protein in MSH1-EPT1 intergenic region.
 GN YHR121W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
 Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 Nhan M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
 Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
 Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII.";
 RL Science 265:2077-2082(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U00059; AAB68870.1; --
 DR PIR; S48963; S48963.
 DR Germonline; 139438; --
 DR SGD; S0001163; YHR121W.
 KW Hypothetical protein.
 SQ SEQUENCE 187 AA; 21313 MW; 0594E464F662CF49 CRC64;
 Query Match 79.1%; Score 34; DB 1; Length 187;
 Best Local Similarity 85.7%; Pred. No. 7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSYVNVQ 7
 DB 81 PSYVNV 87
 RESULT 8
 ZDHC HUMAN STANDARD; PRT; 267 AA.
 ID ZDHC HUMAN
 AC Q36GR4; Q96T09;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zinc finger DHHC domain containing protein 12 (zinc finger protein
 400).
 DE ZDHHC12 OR ZNF400.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Teratocarcinoma;
 RA Isagai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 Ninomiya K., Iwayanagi T.;
 RA "NEDO human cDNA sequencing project.";
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RL


```

RN  [2]
RP  SEQUENCE FROM N.A. (ISOFORM 2).
RC  TISSUE=teratocarcinoma;
RA  Ota T., Nishikawa T., Suzuki Y., Kawai-Rio Y., Hayashi K., Ishii S.,
RA  Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA  Nagahari K., Sugano S., Isogai T.;
RT  "HRI human cDNA sequencing project.";
RL  Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RW  [3]
RP  SEQUENCE FROM N.A. (ISOFORM 1).
RC  TISSUE=Uterus;
RX  MEDLINE=22388257; PubMed=12477932;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan K.C., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA  Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA  Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA  Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length
RT  human and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC  -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC  -!- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=1;
CC  IsoId=Q96GR4-1; Sequence=Displayed;
CC  Name=2;
CC  IsoId=Q96GR4-2; Sequence=VSP_006945;
CC  Note=No experimental confirmation available;
CC  -!- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC  [2]
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AK027430; BAB55104.1; -;
DR  EMBL; AK075332; BAC11553.1; -;
DR  EMBL; BC009280; AAH09280.1; -;
DR  Genew; HGNC:19159; ZDHHC12.
DR  InterPro; IPR001594; Znf DHHC.
DR  Pfam; PF01529; zf-DHHC; 1.
DR  PROSITE; PS50216; ZF-DHHC; 1.
DR  Transmembrane; Zinc-finger; Alternative splicing.
FT  TRANSMEM 10 30 POTENTIAL.
FT  TRANSMEM 44 64 POTENTIAL.
FT  TRANSMEM 141 161 POTENTIAL.
FT  TRANSMEM 179 199 POTENTIAL.
FT  ZN_FING 97 147 DHHC-TYPE.
FT  VARSP LIC 189 240 Missing (in isoform 2).
FT  /FTId=VSP_006945.
SQ  SEQUENCE 267 AA; 30798 MW; C4785D953F1AC95C CRC64;
Query Match 76.7%; Score 33; DB 1; Length 267;
Best Local Similarity 85.7%; Pred. NO. 16;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 PSYVNVQ 7
| | | | |

```

```

Db 59 PGYVNVQ 75
RESULT 9
POL_HTL1A
ID POL_HTL1A STANDARD; PRT; 896 AA.
AC P03362;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [Contains: Reverse transcriptase (EC 2.7.7.49);
DE Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human T-cell leukemia virus type I (strain ATK) (HTLV-I).
OC Viruses; Retroviridae; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11926;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83221647; PubMed=6304725;
RA Seiki M., Hattori S., Hirayama Y., Yoshida M.;
RT "Human adult T-cell leukemia virus: complete nucleotide sequence of
RT the provirus genome integrated in leukemia cell DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3618-3622 (1983).
RN [2]
RP SEQUENCE OF 69-185 FROM N.A.
RX MEDLINE=89210803; PubMed=2468487;
RA Bangham C.R.M., Daenke S., Phillips R.E., Cruickshank J.K.,
RA Bell J.I.;
RT "Enzymatic amplification of exogenous and endogenous retroviral
RT sequences from DNA of patients with tropical spastic paraparesis.";
RL EMBO J. 7:4179-4184 (1988).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC [2]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; J02029; AAA96673.1; -;
DR  EMBL; X14144; CAA32360.1; -;
DR  PIR; A03961; GNLJGH.
DR  HSP; P03355; 1MML.
DR  InterPro; IPR001037; Integrase C.
DR  InterPro; IPR002156; RNaseH.
DR  InterPro; IPR001584; Rve.
DR  InterPro; IPR000477; RVTse.
DR  Pfam; PF00552; integrase; 1.
DR  Pfam; PF00075; rnaseh; 1.
DR  Pfam; PF00665; rve; 1.
DR  Pfam; PF00078; rvt; 1.
DR  Polyprotein; Hydrolase; Endonuclease; Nuclease; Transferase;
KW RNA-directed DNA polymerase.
SQ  SEQUENCE 896 AA; 100141 MW; 113D45D4BD79C65F CRC64;
Query Match 76.7%; Score 33; DB 1; Length 896;
Best Local Similarity 62.5%; Pred. NO. 61;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 PSYVNVQ 8
| | | | |
Db 715 PSYINTDN 722
RESULT 10
POL_HTL1C

```

ID POL HTLIC STANDARD; PRT; 896 AA.
AC P14078;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL polyprotein [Contains: Reverse transcriptase (EC 2.7.7.49);
DE Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human T-cell leukemia virus type I (Caribbean isolate) (HTLV-I).
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11927;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88274338; PubMed=2899128;
RA Malik K.T.A., Even J., Karpas A.;
RT "Molecular cloning and complete nucleotide sequence of an adult T
RT cell leukaemia virus/human T cell leukaemia virus type I
RT (ATLV/HTLV-I) isolate of Caribbean origin: relationship to other
RT members of the ATLV/HTLV-I subgroup.";
RL J. Gen. Virol. 69:1695-1710(1988).
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- PTM: Cleavage sites that yield the mature proteins remain to be
CC determined.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13784; BAA02931.1; --
DR PIR; C28136; GNLCJN.
DR HSSP; P03355; 1MML.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVase.
DR Pfam; PF00552; integrase; 1.
DR Pfam; PF00075; rnaaseH; 1.
DR Pfam; PF00665; rve; 1.
DR Pfam; PF00078; rvt; 1.
DR Polyprotein; Hydrolase; Endonuclease; Nuclease; Transferase;
KW RNA-directed DNA polymerase.
SQ SEQUENCE 896 AA; 100199 MW; 78EF5347EF3E6136 CRC64;
Query Match 76.7%; Score 33; DB 1; Length 896;
Best Local Similarity 62.5%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYVNVQN 8
Db |||:| |
715 PSYINTDN 722
RESULT 11
ID BGAL_LACDE STANDARD; PRT; 1006 AA.
AC P20043;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Beta-galactosidase (EC 3.2.1.23) (Lactase).
GN LACZ.
OS Lactobacillus delbrueckii (subsp. bulgaricus).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1585;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-29.
RC STRAIN=B131;
RX MEDLINE=89123132; PubMed=2492511;
RA Schmidt B.F., Adams R.M., Requaad C., Power S., Mainzer S.E.;
RT "Expression and nucleotide sequence of the Lactobacillus bulgaricus
RT beta-galactosidase gene cloned in Escherichia coli.";
RL J. Bacteriol. 171:625-635(1989).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 11842;
RX MEDLINE=91161509; PubMed=1705929;
RA Leong-Morgenthauer P.M., Zwaalen M.C., Hottinger H.;
RT "Lactose metabolism in Lactobacillus bulgaricus: analysis of the
RT primary structure and expression of the genes involved.";
RL J. Bacteriol. 173:1951-1957(1991).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to family 2 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M23530; AAA25240.1; --
DR EMBL; M55068; AAA25244.1; --
DR PIR; A30093; A30093.
DR HSSP; P00722; IBGL.
DR InterPro; IPR008979; Gal_bind_like.
DR InterPro; IPR006101; Glyco_hydro_2.
DR InterPro; IPR006102; Glyco_hydro_2IG.
DR InterPro; IPR006104; Glyco_hydro_2SB.
DR InterPro; IPR006103; Glyco_hydro_2TIM.
DR InterPro; IPR004200; Glyco_hydro_42C.
DR InterPro; IPR004199; Glyco_hydro_42N.
DR Pfam; PF02930; Bgal_small_C; 1.
DR Pfam; PF02929; Bgal_small_N; 1.
DR Pfam; PF00703; Glyco_hydro_2; 1.
DR Pfam; PF02836; Glyco_hydro_2_C; 1.
DR Pfam; PF02837; Glyco_hydro_2_N; 1.
DR PRINTS; PR00132; GLHYDLASE2.
DR PROSITE; PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE; PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KW Hydrolase; Glycosidase.
FT INIT_MET 0 0
FT ACT_SITE 464 464 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 531 531 NUCLEOPHILE (BY SIMILARITY).
FT CONFLICT 903 903 Y -> YK (IN REF. 2).
SQ SEQUENCE 1006 AA; 113915 MW; B5F81240EB64E769 CRC64;
Query Match 76.7%; Score 33; DB 1; Length 1006;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVNVQ 7
Db | | | | |
100 PQYVNVQ 106
RESULT 12
ID Y435_BUCAI STANDARD; PRT; 319 AA.
AC P57510;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BU435.
GN BU435.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

```

OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RL Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SIMILARITY: STRONG, TO E.COLI YGFZ, ALSO TO H.INFLUENZAE HI0466.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001119; BAB1313.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 319 AA; 37099 MW; 021D7B17DEB52A9E CRC64;

Query Match 74.4%; Score 32; DB 1; Length 319;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 2 PSFISIQN 9

RESULT 13
Y226 MYCPN
ID Y226 MYCPN STANDARD; PRT; 503 AA.
AC P75462;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG226 homolog (F10_orf503).
GN MPN319 OR MP517.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.GENITALIUM MG225.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000051; AAB96165.1; -.
DR PIR; S73843;
DR InterPro; IPR002293; AA/re1.permease1.
DR InterPro; IPR004841; Permease region.
DR Pfam; PF00324; aa.permeases.1.
KW Hypothetical protein; Transmembrane; Complete proteome.

```

```

FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 359 379 POTENTIAL.
FT TRANSMEM 405 425 POTENTIAL.
FT TRANSMEM 443 463 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
SQ SEQUENCE 503 AA; 54960 MW; 4BC1BFDE036985B2 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 503;
Best Local Similarity 72.5%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 343 PSFINEQN 350

RESULT 14
Y85K SSV1
ID Y85K SSV1 STANDARD; PRT; 792 AA.
AC P20210;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Hypothetical 85.7 kDa protein (ORF C-792).
OS Sulfolobus virus-like particle SSV1.
OC Viruses; dsDNA viruses, no RNA stage; Fuselloviridae; Fusellovirus.
OX NCBI_TaxID=244589;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024080; PubMed=1926776;
RA Palm P., Schleper C., Grampp B., Yeats S., McWilliam P., Reiter W.-D.,
RA Zillig W.;
RT "Complete nucleotide sequence of the virus SSV1 of the
RT archaeobacterium Sulfolobus shibatae.";
RL Virology 185:242-250(1991).
CC -!- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN VIRUS ASSEMBLY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X07234; CAA30200.1; -.
DR PIR; S03232; S03232.
KW Hypothetical protein.
SQ SEQUENCE 792 AA; 85657 MW; F4B6426F577396AF CRC64;

Query Match 74.4%; Score 32; DB 1; Length 792;
Best Local Similarity 75.0%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 65 PSYVSVYN 72

RESULT 15
CHS2 RH1OL
ID CHS2 RH1OL STANDARD; PRT; 858 AA.
AC P30595;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)

```

DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
transferase 2).
CHS2.
GN Rhizopus oligosporus.
OS Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
OX NCBI_TaxID=4847;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95036875; PubMed=7765484;
RA Motoyama T., Sudoh M., Horiuchi H., Ohta A., Takagi M.;
RT "Isolation and characterization of two chitin synthase genes of
Rhizopus oligosporus";
RL Biosci. Biotechnol. Biochem. 58:1685-1693(1994).
CC -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-
beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-
glucosaminyl)}(N+1).
CC SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -!- SIMILARITY: Belongs to the chitin synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D10160; BAA01024.1; -.
DR PIR; JC2309;
DR InterPro; IPR004834; Chitin synth.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF01644; Chitin synth; 1.
DR ProDom; PD002998; Chitin synth; 1.
DR Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
SQ SEQUENCE 858 AA; 96791 MW; 5B2DECF5823408C CRC64;

Query Match 74.4%; Score 32; DB 1; Length 858;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||| |||
Db 702 PSYTNVLN 709

RESULT 16
CHS2_EXCODE STANDARD; PRT; 928 AA.
AC P30601; O74210;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
transferase 2) (Class-I chitin synthase 2).
GN CHS2.
OS Exophiala dermatitidis (Wangiella dermatitidis).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Chaetothyriomycetes;
OC Chaetothyriales; Herpotrichiellaceae; mitosporic Herpotrichiellaceae;
OC Exophiala.
OX NCBI_TaxID=5970;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=8656;
RA Zheng L., Szaniszlo P.J.;
RT "Cloning and characterization of WDCHS2, a class I chitin synthase
gene, in Wangiella (Exophiala) dermatitidis";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 259-462 FROM N.A.
RX MEDLINE=92115692; PubMed=1731323;

RA Bowen A.R., Chen-Wu J.L., Momany M., Young R., Szaniszlo P.J.,
RA Robbins P.W.;
RT "Classification of fungal chitin synthases";
RL Proc. Natl. Acad. Sci. U.S.A. 89:519-523(1992).
CC -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-
beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-
glucosaminyl)}(N+1).
CC SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -!- SIMILARITY: Belongs to the chitin synthase family. Subfamily class
I.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF052606; AAC34496.1; -.
DR EMBL; M81906; AAA30335.1; -.
DR PIR; A45189; A45188; Chitin synth.
DR InterPro; IPR004834; Chitin synth.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF01644; Chitin synth; 1.
DR ProDom; PD002998; Chitin synth; 1.
DR Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT CONFLICT 259 260 VV -> KL (IN REF. 2).
FT CONFLICT 462 462 M -> L (IN REF. 2).
SQ SEQUENCE 928 AA; 105712 MW; 001AF053F8873C9B CRC64;

Query Match 74.4%; Score 32; DB 1; Length 928;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||| |||
Db 762 PSYTNVLN 769

RESULT 17
AS10_YEAST STANDARD; PRT; 1146 AA.
ID AS10_YEAST
AC P48361;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ASK10 protein.
GN ASK10 OR YGR097W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=S288C / SEY6210;
RX MEDLINE=97060018; PubMed=8904339;
RA Page N., Sheraton J., Brown J.L., Stewart R.S., Bussey H.;
RT "Identification of ASK10 as a multicopy activator of Skn7p-dependent
transcription of a HIS3 reporter gene";
RL Yeast 12:267-272(1996).
[2]
RN SEQUENCE FROM N.A.
RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Putative activator of SKN7.
CC -!- SIMILARITY: TO YEAST YJDI05C AND YNL047C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

CC -----
DR EMBL; U27209; AAA67368.1; -.
DR EMBL; 272882; CAA97100.1; -.
DR PIR; S64402; S64402.
DR GerOnline; 141409; -.
DR TRANSFAC; T03201; -.
DR SGD; S0003329; ASK10.
DR GO; GO:0006350; P:transcription; IGI.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR DOMAIN 22 26 POLY-GLY.
FT DOMAIN 22 26 POLY-SER.
FT DOMAIN 625 628 POLY-SER.
FT DOMAIN 933 938 POLY-ASN.
FT DOMAIN 958 961 POLY-GLN.
FT DOMAIN 972 975 POLY-SER.
FT CONFLICT 57 57 I -> T (IN REF. 1).
FT CONFLICT 346 346 T -> N (IN REF. 1).
FT CONFLICT 464 464 P -> R (IN REF. 1).
FT CONFLICT 467 467 V -> A (IN REF. 1).
FT CONFLICT 603 603 A -> V (IN REF. 1).
FT CONFLICT 906 906 L -> P (IN REF. 1).
SQ SEQUENCE 1146 AA; 126863 MW; 9E9EF22B0EB496EE CRC64;

Query Match 74.4%; Score 32; DB 1; Length 1146;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYNNVQN 8
Db 784 PQYIHQN 791

RESULT 18
ID RLPA NEIMA STANDARD; PRT; 239 AA.
AC Q9USM7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RlpA-like protein precursor.
GN NMA2219.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=2022556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491".
RL Nature 404:502-506(2000).
CC -!- SIMILARITY: Belongs to the rlpA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

DR EMBL; AL162758; CAB85430.1; -.
DR PIR; F81795; F81795.
DR InterPro; IPR005132; Lipoprotein_13.
DR Pfam; PF03330; Lipoprotein_13; 1.
DR TIGRFAMs; TIGR00413; rlpA; 1.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 239 RLPA-LIKE PROTEIN.
SQ SEQUENCE 239 AA; 26029 MW; D2EDAD47382A275D CRC64;

Query Match 72.1%; Score 31; DB 1; Length 239;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYNNVQN 8
Db 97 PSYVRVTN 104

RESULT 19
ID RLPA NEIMB STANDARD; PRT; 239 AA.
AC Q9KLA0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RlpA-like protein precursor.
GN NMB0267.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58".
RL Science 287:1809-1815(2000).
CC -!- SIMILARITY: Belongs to the rlpA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.)

CC -----
DR EMBL; AE002383; AAF40721.1; -.
DR PIR; D81218; D81218.
DR TIGR; NMB0267; -.
DR InterPro; IPR005132; Lipoprotein_13.
DR Pfam; PF03330; Lipoprotein_13; 1.
DR TIGRFAMs; TIGR00413; rlpA; 1.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 239 RLPA-LIKE PROTEIN.
SQ SEQUENCE 239 AA; 25994 MW; 629EA8EFB7DF693A CRC64;

Query Match 72.1%; Score 31; DB 1; Length 239;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYNNVQN 8

```

Db          97 PSYVRVTN 104
RESULT 20
RLPA_VIBCH
ID  RLPA_VIBCH  STANDARD;  PRT;  263 AA.
AC  Q9KTF4;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  RlpA-like lipoprotein precursor.
GN  VC0948.
OS  Vibrio cholerae.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC  Vibrionaceae; Vibrio.
OX  NCBI_TaxID=666;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=El Tor N16961 / Serotype O1;
RX  MEDLINE=20406833; PubMed=10952301;
RA  Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA  Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA  Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA  Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA  McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA  Salzberg S.B., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA  Fraser C.M.;
RT  "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT  cholerae.";
RL  Nature 406:477-483 (2000).
CC  -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC  (Potential).
CC  -1- POTENTIAL: Belongs to the rlpA family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; A5004177; AAF94110.1; -.
CC  FIRM; G82259; G82259.
CC  TIGR; VC0948; -.
CC  InterPro; IPR005132; Lipoprotein I3.
CC  InterPro; IPR000437; Prok_lipoprot_s.
CC  InterPro; IPR007730; SPOR.
CC  Pfam; PF03330; Lipoprotein_13; 1.
CC  Pfam; PF05036; SPOR; 2.
CC  TIGRfam; TIGR00413; rlpA; 2.
CC  PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC  Hypothetical protein; Membrane; Lipoprotein; Signal;
CC  Complete proteome; Palmitate.
CC  FT  SIGNAL 1 16 POTENTIAL.
CC  FT  CHAIN 17 263 RLPA-LIKE LIPOPROTEIN.
CC  FT  LIPID 17 17 N-Palmitoyl cysteine (Potential).
CC  FT  LIPID 17 17 S-diacylglycerol cysteine (Potential).
CC  SQ  SEQUENCE 263 AA; 29290 MW; 3E5B2E7BD67050EE CRC64;

Query Match 72.1%; Score 31; DB 1; Length 263;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 114 PSYVNVQN 121

RESULT 21
DAPD_BORPE
ID  DAPD_BORPE  STANDARD;  PRT;  273 AA.

```

```

AC  Q9ZEX2;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE  [EC 2.3.1.117] (Tetrahydropicolinate N-succinyltransferase)
DE  (THP succinyltransferase) (Tetrahydropicolinate succinylase).
GN  DAPD OR EPI764 OR BS2183.
OS  Bordetella pertussis, and
OS  Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC  Alcaligenaceae; Bordetella.
OX  NCBI_TaxID=520, 518;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  SPECIES=B.pertussis; STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RC  MEDLINE=20309689; PubMed=10850974;
RX  Fuchs T.M., Schneider B., Krumbach K., Eggeling L., Gross R.;
RT  "Characterization of a Bordetella pertussis diaminopimelate (DAP)
RT  biosynthesis locus identifies dapC, a novel gene coding for an
RT  N-succinyl-L,D-DAP aminotransferase.";
RL  J. Bacteriol. 182:3626-3631 (2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  SPECIES=B.pertussis, and B.bronchiseptica;
RC  STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251, and RB50 / ATCC BAA-588;
RX  MEDLINE=22827954; PubMed=12910271;
RA  Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA  Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA  Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA  Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA  Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA  Feilwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA  Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA  Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA  Sharp S., Simmonds M., Skellon J., Squares R., Squares S., Stevens K.,
RA  Unwin L., Whitehead S., Bartell B.G., Maskell D.J.;
RT  "Comparative analysis of the genome sequences of Bordetella pertussis,
RT  Bordetella parapertussis and Bordetella bronchiseptica.";
RL  Nat. Genet. 35:332-40 (2003).
CC  -1- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC  2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC  oxoheptanedioate.
CC  -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC  semialdehyde; fourth step.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE CYSE/JACA/LPXA/NODL FAMILY OF
CC  ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AJ009834; CA08875.1; -.
CC  EMBL; BX640416; CAE42051.1; -.
CC  EMBL; BX640443; CAE32679.1; -.
CC  HSSP; P56220; 1TDT.
CC  InterPro; IPR005664; DapD.
CC  InterPro; IPR001451; Hexapep_transf.
CC  Pfam; PF00132; hexapep; 5.
CC  TIGRfam; TIGR00965; dapD; 1.
CC  PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
CC  Transfrase; Acyltransferase; Repeat; Lysine biosynthesis;
CC  Diaminopimelate biosynthesis; Complete proteome.
CC  SQ  SEQUENCE 273 AA; 29264 MW; 304456DD34293EF3 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 273;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 PSYVNV 6
DB      126 PSYVNI 131

RESULT 22
DAPD BUCAP
ID DAPD BUCAP STANDARD; PRT; 273 AA.
AC O85290;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE (EC 2.3.1.117) (Tetrahydronicolinate N-succinyltransferase)
DE (THP succinyltransferase) (Tetrahydronicolinate succinylase).
GN DAPD OR BUSG223.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamás I., Klason L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
RN [2]
RP SEQUENCE OF 103-273 FROM N.A.
RX MEDLINE=98353428; PubMed=9688822;
RA Thao M.L., Baumann P.;
RT "Sequence analysis of a DNA fragment from Buchnera aphidicola (Aphid
RT endosymbiont) containing the genes dapD-htrA-ilvH-ftsL-ftsI-
RT murE.";
RL Curr. Microbiol. 37:214-216(1998).
CC -!- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -!- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF014098; AAC67782.1; -.
DB EMBL; AF060492; AAC32330.1; -.
DB HSSP; P56220; 1TDT.
DB InterPro; IPR005664; DapD.
DB InterPro; IPR001451; Hexapep_transf.
DB Pfam; PF00132; hexapep; 5.
DB TIGRFAMs; TIGR00965; dapD; 1.
DB PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis.
SQ SEQUENCE 273 AA; 30927 MW; 58D7AEE35DE356C9 CRC64;
-----
Query Match 72.1%; Score 31; DB 1; Length 273;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
DB      125 PSYVNI 130

RESULT 24
DAPD ECOLI
ID DAPD ECOLI STANDARD; PRT; 274 AA.
AC P03948;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE (EC 2.3.1.117) (Tetrahydronicolinate N-succinyltransferase)
DE (THP succinyltransferase) (Tetrahydronicolinate succinylase).
GN DAPD ECOLI
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S 4074 / Serotype 1;
RX MEDLINE=94224145; PubMed=8170389;
RA Lalonde G., O'Hanley P.D., Stocker B.A., Denich K.;
RT "Characterization of a 3-dehydroquinase gene from Actinobacillus
RT pleuropneumoniae with homology to the eukaryotic genes qa-2 and
RT QUTE.";
RL Mol. Microbiol. 11:273-280(1994).
CC -!- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -!- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X53201; CAB44983.1; -.
DB PIR; S18857; S18857.
DB HSSP; P56220; 3TDT.
DB InterPro; IPR005664; DapD.
DB InterPro; IPR001451; Hexapep_transf.
DB Pfam; PF00132; hexapep; 5.
DB TIGRFAMs; TIGR00965; dapD; 1.
DB PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis.
SQ SEQUENCE 274 AA; 29761 MW; 8AB1CB71FA53455E CRC64;
-----
Query Match 72.1%; Score 31; DB 1; Length 274;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
DB      124 PSYVNI 129

RESULT 24
DAPD ECOLI
ID DAPD ECOLI STANDARD; PRT; 274 AA.
AC P03948;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE (EC 2.3.1.117) (Tetrahydronicolinate N-succinyltransferase)
DE (THP succinyltransferase) (Tetrahydronicolinate succinylase).
GN DAPD ECOLI
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S 4074 / Serotype 1;
RX MEDLINE=94224145; PubMed=8170389;
RA Lalonde G., O'Hanley P.D., Stocker B.A., Denich K.;
RT "Characterization of a 3-dehydroquinase gene from Actinobacillus
RT pleuropneumoniae with homology to the eukaryotic genes qa-2 and
RT QUTE.";
RL Mol. Microbiol. 11:273-280(1994).
CC -!- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -!- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X53201; CAB44983.1; -.
DB PIR; S18857; S18857.
DB HSSP; P56220; 3TDT.
DB InterPro; IPR005664; DapD.
DB InterPro; IPR001451; Hexapep_transf.
DB Pfam; PF00132; hexapep; 5.
DB TIGRFAMs; TIGR00965; dapD; 1.
DB PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis.
SQ SEQUENCE 274 AA; 29761 MW; 8AB1CB71FA53455E CRC64;
-----
Query Match 72.1%; Score 31; DB 1; Length 274;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
DB      125 PSYVNI 130

```

GN DAPD OR B0166 OR SF0156 OR S0159.
OS Escherichia coli, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Escherichia.
[1]
RN NCB1_taxid=562, 623;
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=85054973; PubMed=6094577;
RA Richard C., Richaud F., Martin C., Haziza C., Patte J.-C.;
RT "Regulation of expression and nucleotide sequence of the Escherichia
RT coli dapd gene.";
RL J. Biol. Chem. 259:14824-14828(1984).
[2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RX MEDLINE=94261430; PubMed=8202364;
RA Fujita N., Mori H., Yura T., Ishihama A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 2.4-4.1 min (110,917-193,643 bp) region.";
RL Nucleic Acids Res. 22:1637-1639(1994).
[3]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[4]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federpiet N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lashkari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE OF 1-15 FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RX MEDLINE=94018640; PubMed=8412694;
RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.D.;
RT "The genes of the glutamine synthetase adenylation cascade are not
RT regulated by nitrogen in Escherichia coli.";
RL Mol. Microbiol. 9:443-458(1993).
[6]
RN SEQUENCE OF 1-11.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RX Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
RA Hochstrasser D.F.;
RT Submitted (SEP-1994) to Swiss-Prot.
[7]
RN SEQUENCE OF 1-12.
RC SPECIES=E.coli; STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
[8]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao X., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity

through comparison with genomes of Escherichia coli K12 and O157.";
Nucleic Acids Res. 30:4432-4441(2002).
[9]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NOXL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; K02970; AAA23667.1; -;
DR EMBL; D26562; BAB96742.1; -;
DR EMBL; AE000126; AAC73277.1; -;
DR EMBL; U70214; AAB08595.1; -;
DR EMBL; Z21842; CAA79888.1; -;
DR EMBL; AE015052; AAN41818.1; -;
DR EMBL; AE016978; AAP15699.1; -;
DR PIR; F64740; XNECSD.
DR HSP; P56220; 3TDT.
DR SWISS-2DPAGE; P03948; COLI.
DR EcoGene; EG10207; dapD.
DR InterPro; IPR005664; DapD.
DR InterPro; IPR001451; Hexapep_transf.
DR Pfam; PF00132; hexapep; 4.
DR TIGRfams; TIGR00965; dapD; 1.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis; Complete proteome.
FT CONFLICT 31 V -> D (IN REF. 1).
FT CONFLICT 163 G -> R (IN REF. 1).
FT CONFLICT 177 I -> M (IN REF. 1).
FT CONFLICT 190 V -> L (IN REF. 1).
SQ SEQUENCE 274 AA; 29892 MW; 42D7A38610DD3AF6 CRC64;
Query Match 72.1%; Score 31; DB 1; Length 274;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 PSVVNV 6
Db 125 PSVNI 130
|||||
RESULT 25
DAPD MYCBO STANDARD; PRT; 274 AA.
AC P56220;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE (EC 2.3.1.117) (Tetrahydrodipicolinate N-succinyltransferase)
DE (THP succinyltransferase) (Tetrahydrodipicolinate succinylase).

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD.";
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H₂O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NOXL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32836; AAC23279.1; ALT_INIT.
CC HSP; P56220; 3TDT.
CC TIGR; H11634; -.
CC InterPro; IPR005664; DapD.
CC InterPro; IPR001451; Hexapep_transf.
CC Pfam; PF00132; hexapep; 5.
CC TIGRFAMs; TIGR00965; dapD; 1.
CC PROSITE; PS00101; HEXAPEP TRANSFERASES; 1.
CC Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
CC Diaminopimelate biosynthesis; Complete proteome.
CC KW Diaminopimelate biosynthesis; Complete proteome.
CC SEQUENCE 275 AA; 29723 MW; 6058EC21B4283498 CRC64;
CC -----
CC Query Match 72.1%; Score 31; DB 1; Length 275;
CC Best Local Similarity 83.3%; Pred. No. 44;
CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 PSYVNV 6
CC | | | | |
CC Db 125 PSYVNI 130
CC -----
CC RESULT 27
CC Y360 HAEIN
CC ID Y360 HAEIN STANDARD; PRT; 282 AA.
CC AC P44661;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Probable iron transport system membrane protein HI0360.
CC GN HI0360.
CC OS Haemophilus influenzae.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC OC Pasteurellaceae; Haemophilus.
CC OX NCBI_TaxID=727;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=RD / KW20 / ATCC 51907;
CC RX MEDLINE=95350630; PubMed=7542800;
CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
CC RA Karvavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
CC RA McKenney K., Sutton R., Fitzhugh W., Fields C.A., Gocayne J.D.,
CC RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
CC RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
CC RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
CC RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
CC RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
CC RA Venter J.C.;
CC "Whole-genome random sequencing and assembly of Haemophilus influenzae
CC RD.";
CC RL Science 269:496-512(1995).
CC -----

CC -1- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC HI0359/HI0360/HI0361/HI0362 FOR IRON (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32720; AAC22019.1; -.
CC PIR; E64063; E64063.
CC TIGR; HI0360; -.
CC InterPro; IPR001626; ABC_transp3.
CC Pfam; PF00950; ABC-3; 1.
CC Hypothetical protein; Transport; Transmembrane; Inner membrane;
CC KW Iron transport; Complete proteome.
CC FT TRANSMEM 17 37 POTENTIAL.
CC FT TRANSMEM 63 83 POTENTIAL.
CC FT TRANSMEM 93 113 POTENTIAL.
CC FT TRANSMEM 140 160 POTENTIAL.
CC FT TRANSMEM 164 184 POTENTIAL.
CC FT TRANSMEM 186 206 POTENTIAL.
CC FT TRANSMEM 223 243 POTENTIAL.
CC FT TRANSMEM 245 265 POTENTIAL.
CC SQ SEQUENCE 282 AA; 30293 MW; 5EB442874214748C CRC64;
CC -----
CC Query Match 72.1%; Score 31; DB 1; Length 282;
CC Best Local Similarity 75.0%; Pred. No. 45;
CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 PSYVNVQN 8
CC | | | | |
CC Db 114 PTAIVNVQN 121
CC -----
CC RESULT 28
CC RLPA PSEAE
CC ID RLPA PSEAE STANDARD; PRT; 342 AA.
CC AC Q3XV6;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Rlpa-like lipoprotein precursor.
CC GN RLPA OR PA4000.
CC OS Pseudomonas aeruginosa.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC OC Pseudomonadaceae; Pseudomonas.
CC OX NCBI_TaxID=287;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 15692 / PA01;
CC RA Gagnon L.A., Castro-Urbina I.M., Liao X., Hancock R.E.W., Clarke A.J.,
CC RA Huletsky A.;
CC "Cloning and characterization of PPP5 of Pseudomonas aeruginosa.";
CC RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 15692 / PA01;
CC RX MEDLINE=20437337; PubMed=10984043;
CC RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
CC RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
CC RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
CC RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
CC RA Smith K.J., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
CC RA Reizer J., Sauer M.H., Hancock R.E.W., Lory S., Olson M.V.;
CC "Complete genome sequence of Pseudomonas aeruginosa PA01, an
CC opportunistic pathogen.";

```

RL Nature 406:959-964(2000).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -!- SIMILARITY: Belongs to the rlpA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF147448; AAD32233.1; -.
DR EMBL; AE004817; AGA07387.1; -.
DR PIR; D83146; D83146.
DR InterPro; IPR005132; Lipoprotein_13.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR InterPro; IPR007730; SPOR.
DR Pfam; PF03330; Lipoprotein_13; 1.
DR Pfam; PF05036; SPOR; 2.
DR TIGRFAMs; TIGR00413; rlpA; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Hypothetical protein; Membrane; Lipoprotein; Signal;
KW Complete proteome; Palmitate.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 342 RLPA-LIKE LIPOPROTEIN
FT LIPID 27 27 N-palmitoyl cysteine (potential).
FT LIPID 27 27 S-diacylglycerol cysteine (potential).
SQ SEQUENCE 342 AA; 36482 MW; 18D08C614E23E24D CRC64;

Query Match 72.1%; Score 31; DB 1; Length 342;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYNNVQN 8
Db 137 PSYNNVQN 144

RESULT 29
FMA2 PORGI STANDARD; PRT; 348 AA.
AC Q51822; Q51823; Q51824;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major fimbrial subunit protein, type II precursor (Fimbrillin)
DE (Fimbrillin).
GN FIMA.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HW24D1, OMZ314, and OMZ409;
RX MEDLINE=94071950; PubMed=7902712;
RA Fujiwara T., Morishima S., Takahashi I., Hanada S.;
RT "Molecular cloning and sequencing of the fimbrillin gene of
RT Porphyromonas gingivalis strains and characterization of recombinant
RT proteins.";
RL Biochem. Biophys. Res. Commun. 197:241-247(1993).
RN [2]
RP FUNCTION, AND CLASSIFICATION INTO TYPES.
RX MEDLINE=21614934; PubMed=11748193;
RA Nakagawa I., Amano A., Kubonishi M., Nakamura T., Kawabata S.,
RA Hanada S.;
RT "Functional differences among FimA variants of Porphyromonas
RT gingivalis and their effects on adhesion to and invasion of human
RT epithelial cells.";
RL Infect. Immun. 70:277-285(2002).
CC -!- FUNCTION: Fimbrillin is the structural subunit of the fimbriae,

```

```

CC that are filamentous appendages on the cell surface. Fimbriae of
CC P.gingivalis are recognized as a major virulence factor as they
CC mediate cell adhesion and play an important role in invasion of
CC periodontal tissues.
CC -!- SUBCELLULAR LOCATION: Fimbria.
CC -!- SIMILARITY: Belongs to the P.gingivalis fimbrillin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17797; BAA04623.1; -.
DR EMBL; D17798; BAA04624.1; -.
DR EMBL; D17799; BAA04625.1; -.
DR PIR; JN0916; JN0916.
DR PIR; JN0918; JN0918.
DR InterPro; IPR008110; Fimbrillin.
DR PRINTS; PR01737; FIMBRILLIN.
KW Fimbria; Virulence.
FT PROPEP 1 10 BY SIMILARITY.
FT CHAIN 11 348 MAJOR FIMBRIAL SUBUNIT PROTEIN, TYPE II.
FT VARIANT 54 68 EMKLA -> AMELV (IN STRAIN OMZ409).
FT VARIANT 90 90 A -> T (IN STRAIN OMZ409).
FT VARIANT 100 100 E -> D (IN STRAIN OMZ409).
SQ SEQUENCE 348 AA; 38089 MW; 478F8C4E2A63EE2F CRC64;

Query Match 72.1%; Score 31; DB 1; Length 348;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYNNVQN 8
Db 147 PSYNNVQN 154

RESULT 30
RLPA YERPE STANDARD; PRT; 360 AA.
AC Q8ZDG6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Rare lipoprotein A precursor.
GN RLPA OR YPO2602 OR Y1176.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,

```

RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- SIMILARITY: Belongs to the rlpA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ414153; CAC92845.1; -;
DR EMBL; AE013721; AAM84753.1; -;
DR PIR; AF0317; AF0317.
DR InterPro; IPR005132; Lipoprotein 13.
DR InterPro; IPR000437; Prok lipoprot_s.
DR InterPro; IPR007730; SPOR_1.
DR Pfam; PF03330; Lipoprotein_13; 1.
DR Pfam; PF05036; SPOR; 2.
DR TIGRFAMs; TIGR00413; rlpA; 1.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Membrane; Lipoprotein; Signal; Complete proteome; Palmitate.
FT SIGNAL 1 17 BY SIMILARITY.
FT CHAIN 18 360 RARE LIPOPROTEIN A.
FT LIPID 18 18 N-Palmitoyl cysteine.
FT LIPID 18 18 S-diacylglycerol cysteine.
SQ SEQUENCE 360 AA; 38018 MW; 868BC3E44AED401 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 360;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 0;

QY 1 PSYVNVQN 8
Db 113 PSYVRVTN 120
|||||

RESULT 31
NQ04 THETH STANDARD; PRT; 409 AA.
AC Q5620;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NADH-quinone oxidoreductase chain 4 (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain 4) (NDH-1, chain 4).
GN NQ04.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=97172490; PubMed=9020134;
RA Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
RT thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
RT sequence of the gene cluster and thermostable properties of the
RT expressed NQ02 subunit.";
RL J. Biol. Chem. 272:4201-4211(1997).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur [Fe-S] centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be menaquinone. Couples the redox reaction to proton
CC translocation [for every two electrons transferred, four hydrogen
CC ions are translocated across the cytoplasmic membrane], and thus
CC conserves the redox energy in a proton gradient.
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.

CC -1- SUBUNIT: Composed of 14 different subunits.
CC -1- SIMILARITY: Belongs to the complex I 49 kDa subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52917; AAA97941.1; -;
DR PIR; T11901; T11901.
DR InterPro; IPR001135; Oxidored 49kDa.
DR Pfam; PF00346; complex1_49kDa_1.
DR PROSITE; PS00535; COMPLEX1_49K; 1.
KW Oxidoreductase; NAD; Quinone.
SQ SEQUENCE 409 AA; 46371 MW; 161AA0C796D62ED3 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 409;
Best Local Similarity 62.5%; Pred. No. 67;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 373 PSFVNLQS 380
|||||

RESULT 32
YUQP_CABEL STANDARD; PRT; 453 AA.
ID YUQP_CABEL
AC Q19895;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 52.9 kDa protein F28H7.8 in chromosome V.
GN F28H7.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RA Berks M.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 CRAL-TRIO domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z72508; CAA96639.1; -;
DR PIR; T21528; T21528.
DR WormPep; F28H7.8; CE05757.
DR InterPro; IPR001251; CRAL_TRIO_C.
DR InterPro; IPR008273; CRAL_TRIO_N.
DR Pfam; PF00650; CRAL_TRIO; 1.
DR SMART; SM00516; SEC14; 1.
DR PROSITE; PS0191; CRAL_TRIO; 1.
KW Hypothetical protein.
FT DOMAIN 81 257 CRAL-TRIO.
SQ SEQUENCE 453 AA; 52926 MW; CC07AF08D50FDE79 CRC64;

Query Match 72.1%; Score 31; DB 1; Length 453;
Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6

```

Db      203 PSYINV 208
||||:|
RESULT 33
CHSX_USTMA
ID      CHSX_USTMA      STANDARD;      PRT;      760 AA.
AC      Q99126;
DT      15-JUL-1999 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE      transferase 1).
GN      CHS1.
OS      Ustilago maydis (Smut fungus).
OC      Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC      Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX      NCBI_TaxID=5270;
RN      [1]
SEQUENCE FROM N.A.
RC      STRAIN=RK32 / A2B3; PubMed=8932711;
RX      MEDLINE=97086517; PubMed=8932711;
RA      Xocoostle-Cazares B., Leon-Ramirez C., Ruiz-Herrera J.;
RT      "Two chitin synthase genes from Ustilago maydis.";
RL      Microbiology 142:377-387(1996).
CC      -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC      -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + (1,4)-(N-acetyl-
CC      beta-D-glucosaminyl) (N) = UDP + (1,4)-(N-acetyl-beta-D-
CC      glucosaminyl) (N+1).
CC      -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC      -!- SIMILARITY: Belongs to the chitin synthase family.
CC      [1]
SEQUENCE FROM N.A.
RC      STRAIN=RK32 / A2B3; PubMed=8932711;
RX      MEDLINE=97086517; PubMed=8932711;
RA      Xocoostle-Cazares B., Leon-Ramirez C., Ruiz-Herrera J.;
RT      "Two chitin synthase genes from Ustilago maydis.";
RL      Microbiology 142:377-387(1996).
CC      -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC      -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + (1,4)-(N-acetyl-
CC      beta-D-glucosaminyl) (N) = UDP + (1,4)-(N-acetyl-beta-D-
CC      glucosaminyl) (N+1).
CC      -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC      -!- SIMILARITY: Belongs to the chitin synthase family.
CC      [1]
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X87748; CAA61027.1; -.
PIR; S55520; S55520.
InterPro; IPR004834; Chitin synth.
Pfam; PF01644; Chitin synth; 1.
ProDom; PD002998; Chitin synth; 1.
Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
Multigene family.
FT TRANSFER 302 322 POTENTIAL.
FT TRANSFER 385 405 POTENTIAL.
FT TRANSFER 526 546 POTENTIAL.
FT TRANSFER 564 584 POTENTIAL.
FT TRANSFER 602 622 POTENTIAL.
FT TRANSFER 644 664 POTENTIAL.
FT TRANSFER 673 693 POTENTIAL.
FT TRANSFER 778 798 POTENTIAL.
FT TRANSFER 816 836 POTENTIAL.
FT TRANSFER 179 179 E -> EDE (IN REF. 2).
FT TRANSFER 199 200 RV -> HI (IN REF. 2).
FT TRANSFER 352 352 A -> G (IN REF. 2).
SQ SEQUENCE 760 AA; 85181 MW; 9377000F57410993 CRC64;
Query Match 72.1%; Score 31; DB 1; Length 760;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYINVQVN 8
|:|:|:|
Db 748 PTYINLN 755
|:|:|:|
RESULT 34
CHS1_PHYBL
ID      CHS1_PHYBL      STANDARD;      PRT;      841 AA.
AC      P87073;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1999 (Rel. 38, Last annotation update)
DE      Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE      transferase 1) (Class-II chitin synthase 1).
GN      CHS1.
OS      Phycomyces blakesleeanus.
OC      Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC      Phycomyces.
OX      NCBI_TaxID=4937;
RN      [1]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=NRRL 1555;
RA      Miyazaki A., Ootaki T.;
RN      Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
SEQUENCE OF 172-370 FROM N.A.
RP      STRAIN=NRRL 1555;
RX      MEDLINE=94063507; PubMed=8244024;
RA      Miyazaki A., Momany M., Szanislo P.J., Jayaram M., Ootaki T.;
RT      "Chitin synthase-encoding gene(s) of the Zygomycete fungus Phycomyces
RT      blakesleeanus.";
RL      Gene 134:129-134(1993).
CC      -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC      -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + (1,4)-(N-acetyl-
CC      beta-D-glucosaminyl) (N) = UDP + (1,4)-(N-acetyl-beta-D-
CC      glucosaminyl) (N+1).
CC      -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC      -!- SIMILARITY: Belongs to the chitin synthase family.
CC      [1]
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AB003043; BAA19857.1; -.
PIR; JTO767; JTO767.
InterPro; IPR004834; Chitin synth.
DR      InterPro; IPR001173; GlycoTrans_2.
Pfam; PF01644; Chitin synth; 1.
ProDom; PD002998; Chitin synth; 1.
Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
Multigene family.
FT TRANSFER 302 322 POTENTIAL.
FT TRANSFER 385 405 POTENTIAL.
FT TRANSFER 526 546 POTENTIAL.
FT TRANSFER 564 584 POTENTIAL.
FT TRANSFER 602 622 POTENTIAL.
FT TRANSFER 644 664 POTENTIAL.
FT TRANSFER 673 693 POTENTIAL.
FT TRANSFER 778 798 POTENTIAL.
FT TRANSFER 816 836 POTENTIAL.
FT TRANSFER 179 179 E -> EDE (IN REF. 2).
FT TRANSFER 199 200 RV -> HI (IN REF. 2).
FT TRANSFER 352 352 A -> G (IN REF. 2).
SQ SEQUENCE 841 AA; 95226 MW; 71CD6C09ACB6B8B CRC64;
Query Match 72.1%; Score 31; DB 1; Length 841;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYINVQVN 8
|:|:|:|
Db 681 PSYTNILN 688
|:|:|:|
RESULT 35
MTP_HUMAN
ID      MTP_HUMAN      STANDARD;      PRT;      894 AA.
AC      P55157;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Microsomal triglyceride transfer protein, large subunit precursor.
DE      transferase 1) (Class-II chitin synthase 1).
GN      MTP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
SEQUENCE FROM N.A.

```

RC TISSUE=Small intestine;
 RX MEDLINE=94154686; PubMed=8111381;
 RA Shoulders C.C., Brett D.J., Bayliss J.D., Narcisi T.M.E.,
 RA Jarmuz A., Grantham T.T., Leoni P.R.D., Bhattacharya S.,
 RA Pease R.J., Cullen P.M., Levi S., Byfield P.G.H., Furkiss P.,
 RA Scott J.;
 RT "Abetalipoproteinemia is caused by defects of the gene encoding the
 RT 97 kDa subunit of a microsomal triglyceride transfer protein.";
 RL Hum. Mol. Genet. 2:2109-2116(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93368660; PubMed=8361539;
 RA Sharp D., Blinderman L., Combs K.A., Kienzie B., Ricci B.,
 RA Wager-Smith K., Gil C.M., Turck C.W., Bouma M.-E., Rader D.J.,
 RA Aggerbeck L.P., Gregg R.E., Gordon D.A., Wetterau J.R.;
 RT "Cloning and gene defects in microsomal triglyceride transfer protein
 RT associated with abetalipoproteinemia.";
 RL Nature 365:65-69(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94325268; PubMed=7545943;
 RA Sharp D., Ricci B., Kienzie B., Lin M.C., Wetterau J.R.;
 RT "Human microsomal triglyceride transfer protein large subunit gene
 RT structure.";
 RL Biochemistry 33:9057-9061(1994).
 RN [4]
 RP MUTAGENESIS OF CVS-878.
 RX MEDLINE=96065017; PubMed=9533758;
 RA Narcisi T.M.E., Shoulders C.C., Chester S.A., Read J., Brett D.J.,
 RA Harrison G.B., Grantham T.T., Fox M.F., Povey S., de Bruin T.W.A.,
 RA Erkelens D.W., Muller D.P.R., Lloyd J.K., Scott J.;
 RT "Mutations of the microsomal triglyceride-transfer-protein gene in
 RT abetalipoproteinemia.";
 RL Am. J. Hum. Genet. 57:1298-1310(1995).
 RN [5]
 RP SIMILARITY TO VITELLOGENINS.
 RX MEDLINE=95393146; PubMed=7664034;
 RA Shoulders C.C., Narcisi T.M.E., Read J., Chester S.A., Brett D.J.,
 RA Scott J., Anderson T.A., Levitt D.G., Banaszak L.J.;
 RT "The abetalipoproteinemia gene is a member of the vitellogenin family
 RT and encodes an alpha-helical domain.";
 RL Nat. Struct. Biol. 1:285-286(1994).
 RN [6]
 RP VARIANT ABL HIS-540, VARIANTS GLN-237 AND ALA-384, AND MUTAGENESIS OF
 RP ARG-540.
 RX MEDLINE=97094705; PubMed=8939939;
 RA Rehberg E.F., Samson-Bouma M.-E., Kienzie B., Blinderman L., Jamil H.,
 RA Wetterau J.R., Aggerbeck L.P., Gordon D.A.;
 RT "A novel abetalipoproteinemia genotype. Identification of a missense
 RT mutation in the 97-kDa subunit of the microsomal triglyceride
 RT transfer protein that prevents complex formation with protein
 RT disulfide isomerase.";
 RL J. Biol. Chem. 271:29945-29952(1996).
 RN [7]
 RP VARIANTS ABL HIS-540; ILE-590 AND GLU-746.
 RX MEDLINE=20146221; PubMed=10679949;
 RA Wang J., Hegele R.A.;
 RT "Microsomal triglyceride transfer protein (MTP) gene mutations in
 RT Canadian subjects with abetalipoproteinemia.";
 RL Hum. Mutat. 15:294-295(2000).
 RN [8]
 RP VARIANT ABL TYR-780.
 RX MEDLINE=20405694; PubMed=10946006;
 RA Chashi K., Ishibashi S., Oeuga J., Tozawa R., Harada K., Yahagi N.,
 RA Shinoiri F., Iizuka Y., Tamura Y., Nagai R., Illingworth D.R.,
 RA Gotoda T., Yamada N.;
 RT "Novel mutations in the microsomal triglyceride transfer protein gene
 RT causing abetalipoproteinemia.";
 RL J. Lipid Res. 41:1199-1204(2000).
 RN [9]
 RP VARIANTS HIS-95; THR-128; GLU-244 AND GLN-297.
 RX MEDLINE=21652946; PubMed=11792722;

RA Ledmyr H., Karpe F., Lundahl B., McKinnon M., Skoglund-Andersson C.,
 RA Ehrenborg E.;
 RT "Variants of the microsomal triglyceride transfer protein gene are
 RT associated with plasma cholesterol levels and body mass index.";
 RL J. Lipid Res. 43:51-58(2002).
 CC -1- FUNCTION: CATALYZES THE TRANSPORT OF TRIGLYCERIDE, CHOLESTERYL
 CC ESTER, AND PHOSPHOLIPID BETWEEN PHOSPHOLIPID SURFACES. REQUIRED
 CC FOR THE SECRETION OF PLASMA LIPOPROTEINS THAT CONTAIN
 CC APOLIPOPROTEIN B.
 CC -1- SUBUNIT: HETERODIMER COMPOSED OF MTP AND OF PROTEIN DISULFIDE
 CC ISOMERASE (PDI).
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.
 CC -1- TISSUE SPECIFICITY: LIVER AND SMALL INTESTINE. ALSO FOUND IN
 CC OVARY, TESTIS AND KIDNEY.
 CC -1- DISEASE: Defects in mtp are the cause of abetalipoproteinemia
 CC (ABL), an autosomal recessive disorder of lipoprotein metabolism.
 CC Affected individuals produce virtually no circulating
 CC apolipoprotein B-containing lipoproteins (chylomicrons, VLDL,
 CC LDL, lipoprotein(A)). Malabsorption of the antioxidant vitamin E
 CC occurs, leading to spinocerebellar and retinal degeneration.
 CC -1- SIMILARITY: TO VITELLOGENINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X75500; CAA53217.1; -;
 CC EMBL; X59657; CAA42200.1; -;
 CC EMBL; X83013; CAA58142.1; -;
 CC EMBL; X83014; CAA58142.1; JOINED.
 CC EMBL; X83015; CAA58142.1; JOINED.
 CC EMBL; X83016; CAA58142.1; JOINED.
 CC EMBL; X83017; CAA58142.1; JOINED.
 CC EMBL; X83018; CAA58142.1; JOINED.
 CC EMBL; X83019; CAA58142.1; JOINED.
 CC EMBL; X83020; CAA58142.1; JOINED.
 CC EMBL; X83021; CAA58142.1; JOINED.
 CC EMBL; X83022; CAA58142.1; JOINED.
 CC EMBL; X83023; CAA58142.1; JOINED.
 CC EMBL; X83024; CAA58142.1; JOINED.
 CC EMBL; X83025; CAA58142.1; JOINED.
 CC EMBL; X83026; CAA58142.1; JOINED.
 CC EMBL; X83027; CAA58142.1; JOINED.
 CC EMBL; X83028; CAA58142.1; JOINED.
 CC EMBL; X83029; CAA58142.1; JOINED.
 CC EMBL; X83030; CAA58142.1; JOINED.
 CC F1R; I38047; I38047.
 CC Gene; HGNC:7467; MTP.
 CC MIM; 157147; -;
 CC MIM; 200100; -;
 CC GO; GO:0005488; F-binding; TAS.
 CC GO; GO:0006629; P:lipid metabolism; TAS.
 CC InterPro; IPR001747; Lipid_transprt_N.
 CC Pfam; PF01347; Vitellogenin_N; 1.
 CC SMART; SM00638; LpD_N; 1.
 CC Disease mutation.
 CC SIGNAL 1 18 POTENTIAL.
 CC CHAIN 19 894 MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN,
 CC LARGE SUBUNIT.
 CC VARIANT 95 95 Q -> H.
 CC VARIANT 128 128 /FTid=VAR_014016.
 CC VARIANT 244 244 I -> T.
 CC VARIANT 297 297 /FTid=VAR_014017.
 CC VARIANT 297 297 Q -> E.
 CC VARIANT 297 297 H -> Q.
 CC VARIANT 354 354 /FTid=VAR_010640.
 CC VARIANT 354 354 E -> Q (IN dbSNP:129333).

```

FT  VARIANT      384      384      /FTid=VAR_014916.
FT  FT  -> A.
FT  FT  /FTid=VAR_010641.
FT  FT  R -> H (IN ABL; LOSS OF ACTIVITY) .
FT  FT  /FTid=VAR_010642.
FT  FT  S -> I (IN ABL) .
FT  FT  /FTid=VAR_010643.
FT  FT  G -> E (IN ABL) .
FT  FT  /FTid=VAR_010644.
FT  FT  N -> Y (IN ABL; LOSS OF ACTIVITY) .
FT  FT  /FTid=VAR_014019.
FT  FT  R->K: NO CHANGE OF ACTIVITY.
FT  FT  C->S: LOSS OF ACTIVITY.
FT  FT  F -> L (IN REF. 2).
SQ  SEQUENCE      894 AA; 99351 MW; B20260C136BDAB9F CRC64;

Query Match      72.1%; Score 31; DB 1; Length 894;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY  1 PSYVNVQN 8
    |||::||
Db  552 PSYMDVXN 559

RESULT 36
MTP_MOUSE
ID  MTP_MOUSE      STANDARD;      PRT;      894 AA.
AC  O08601;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  Microsomal triglyceride transfer protein, large subunit precursor.
GN  MTP OR MTPP.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver, and Small intestine;
RX  MEDLINE=96301412; PubMed=8660984;
RT  Nakamura M., Chang B.H., Hoocheveen R., Li W.H., Chan L.;
RT  "Mouse microsomal triglyceride transfer protein large subunit: cDNA
RT  cloning, tissue-specific expression and chromosomal localization.";
RL  Genomics 33:313-316(1996).
CC  -1- FUNCTION: CATALYZES THE TRANSPORT OF TRIGLYCERIDE, CHOLESTERYL
CC  ESTER, AND PHOSPHOLIPID BETWEEN PHOSPHOLIPID SURFACES. REQUIRED
CC  FOR THE SECRETION OF PLASMA LIPOPROTEINS THAT CONTAIN
CC  APOLIPOPROTEIN B.
CC  -1- SUBUNIT: HETERODIMER COMPOSED OF MTP AND OF PROTEIN DISULFIDE
CC  ISOMERASE (PDI).
CC  -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.
CC  -1- SIMILARITY: TO VITELLOGENINS.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
-----
EMBL: L47970; AAB51431.1; -.
DR  MGD; MGI:106926; Mtp.
DR  InterPro; IPR001747; Lipid transprt_N.
DR  Pfam; PF01347; Vitellogenin_N; 1.
DR  SMART; SM00638; LPD N; 1.
KW  Endoplasmic reticulum; Lipid-binding; Transport; Signal.
FT  SIGNAL      1      18      POTENTIAL.
FT  CHAIN      19      894      MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN,
                                LARGE SUBUNIT.
SQ  SEQUENCE      894 AA; 99141 MW; 616DGB284CA471555 CRC64;

```

```

Query Match      72.1%; Score 31; DB 1; Length 894;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNQVN 8
DB      551 PSYMDVKN 558
      ||||:|:|
      ||||:|:|

RESULT 37
MTP_MESAU
ID_MTP_MESAU      STANDARD;      PRT;      895 AA.
AC      P55158;
DT      01-OCT-1996 (Rel. 34, Créated)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Microsomal triglyceride transfer protein, large subunit precursor.
GN      MTP.
OS      Mesocricetus auratus (Golden hamster).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC      Mesocricetus.
NCBI_TaxID=10036;
[1]
SEQUENCE FROM N.A.
RC      TISSUE=Intestine;
RX      MEDLINE=95050734; PubMed=7961879;
RA      Lin M.C., Arbeny C., Bergquist K., KiENZle B., Gordon D.A.,
RA      Wetterau J.R.;
RT      "Cloning and regulation of hamster microsomal triglyceride transfer
RT      protein. The regulation is independent from that of other hepatic and
RT      intestinal proteins which participate in the transport of fatty acids
RT      and triglycerides.";
RL      J. Biol. Chem. 269:29138-29145(1994).
CC      -1- FUNCTION: CATALYZES THE TRANSPORT OF TRIGLYCERIDE, CHOLESTERYL
CC      ESTER, AND PHOSPHOLIPID BETWEEN PHOSPHOLIPID SURFACES. REQUIRED
CC      FOR THE SECRETION OF PLASMA LIPOPROTEINS THAT CONTAIN
CC      APOLIPOPROTEIN B.
CC      -1- SUBUNIT: HETERODIMER COMPOSED OF MTP AND OF PROTEIN DISULFIDE
CC      ISOMERASE (PDI).
CC      -1- SUBCELLULAR LOCATION: Endoplasmic reticulum.
CC      -1- SIMILARITY: TO VITELLOGENINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; U14995; AAAS3143.1; -.
DR      PIR; A55413; A55413.
DR      InterPro; IPR001747; Lipid transprt_N.
DR      Pfam; PF01347; Vitellogenin_N; 1.
DR      SMART; SM00638; LPD_N; 1.
KW      Endoplasmic reticulum; Lipid-binding; Transport; Signal.
FT      SIGNAL      1      18      POTENTIAL.
FT      CHAIN      19      895      MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN,
FT      LARGE SUBUNIT.
SQ      SEQUENCE      895 AA; 99388 MW; 841062179548059D CRC64;

Query Match      72.1%; Score 31; DB 1; Length 895;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNQVN 8
DB      552 PSYMDVKN 559
      ||||:|:|
      ||||:|:|

RESULT 38

```

```

CHSA_AMPQU
ID CHSA_AMPQU STANDARD; PRT; 910 AA.
AC Q12564;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chitin synthase A (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE transferase A) (Class-I chitin synthase A).
DE CHSA.
GN Ampelomyces quisqualis.
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Leptosphaeriaceae; mitosporic Leptosphaeriaceae;
OC Ampelomyces
OX NCBI_TaxID=50730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AQ10;
RX MEDLINE=96186963; PubMed=8626074;
RA Weiss N., Szejnberg A., Yarden O.;
RT "The chsa gene, encoding a class-I chitin synthase from Ampelomyces
RT quisqualis.";
RL Gene 168:99-102(1996).
CC -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-
CC beta-D-glucosaminyl)} (N) = UDP + {(1,4)-(N-acetyl-beta-D-
CC glucosaminyl)} (N+1).
CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -!- SIMILARITY: Belongs to the chitin synthase family. Subfamily class
CC I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86802; CAA60497.1; -.
DR F01R; JC4609; JC4609.
DR InterPro; IPR004834; Chitin synth.
DR Pfam; PF01644; Chitin synth. 1.
DR ProDom; PD002998; Chitin synth; 1.
DR Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
DR Multigene family.
FT TRANSFERRASE 366 386 POTENTIAL.
FT TRANSFERRASE 448 468 POTENTIAL.
FT TRANSFERRASE 583 603 POTENTIAL.
FT TRANSFERRASE 620 640 POTENTIAL.
FT TRANSFERRASE 655 675 POTENTIAL.
FT TRANSFERRASE 701 721 POTENTIAL.
FT TRANSFERRASE 730 750 POTENTIAL.
FT TRANSFERRASE 828 848 POTENTIAL.
FT TRANSFERRASE 876 896 POTENTIAL.
SQ SEQUENCE 910 AA; 103012 MW; F8F1DB135F2138F9 CRC64;
Query Match 72.1%; Score 31; DB 1; Length 910;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYVVNVN 8
|:|:|
DB 738 PTVINILN 745
RESULT 39
WRN_HUMAN
ID WRN_HUMAN STANDARD; PRT; 1432 AA.
AC Q14191;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Werner syndrome helicase.
DE WRN OR RECQL2 OR RECQ3.
DE Homo sapiens (Human).
DE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DE NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96181115; PubMed=8602509;
RA Yu C.-E., Oshima J., Fu Y.-H., Wijisman E.M., Hisama F., Alisch R.,
RA Matthews S., Nakura J., Miki T., Ouais S., Martin G.M., Mulligan J.,
RA Schellenberg G.D.;
RT "Positional cloning of the Werner's syndrome gene.";
RL Science 272:258-262(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Paepker B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA Mulligan J., Galas D., Fu Y.-H.;
RT "Genomic structure of the human Werner's gene and cloning of the
RT mouse homolog.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=98284027; PubMed=9618508;
RA Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT "Nucleolar localization of the Werner syndrome protein in human
RT cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RN [4]
RP REPEATS.
RX MEDLINE=99160561; PubMed=10049920;
RA Kusano K., Berres M.B., Engels W.R.;
RT "Evolution of the RECQ family of helicases: a Drosophila homolog,
RT Dmblm, is similar to the human Bloom syndrome gene.";
RL Genetics 151:1027-1039(1999).
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE=99235545; PubMed=10220139;
RA Moser M.J., Oshima J., Monnat R.J. Jr.;
RT "WRN mutations in Werner syndrome.";
RL Hum. Mutat. 13:271-279(1999).
RN [6]
RP VARIANT ARG-1367.
RX MEDLINE=97173161; PubMed=9021029;
RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
RA Ikegami H., Higaki J., Edland S.D., Martin G.M., Ogihara T.;
RT "Association of a polymorphic variant of the Werner helicase gene with
RT myocardial infarction in a Japanese population.";
RL Am. J. Med. Genet. 68:494-498(1997).
RN [7]
RP ERRATUM.
RA Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
RA Ikegami H., Higaki J., Edland S.D., Martin G.M., Ogihara T.;
RL Am. J. Med. Genet. 70:103-103(1997).
RN [8]
RP VARIANTS ILE-387 AND LEU-1074.
RX MEDLINE=98111850; PubMed=9450180;
RA Meiseltzer C., Ruppitsch W., Weirich-Schwaiger H., Weirich H.G.,
RA Jabkowsky J., Klein G., Schweiger M., Hirsch-Kauffmann M.;
RT "Werner syndrome: characterization of mutations in the WRN gene in an
RT affected family.";
RL Eur. J. Hum. Genet. 5:364-370(1997).
RN [9]
RP VARIANT ILE-387.
RA Vidal V., Bay J.-O., Champomier F., Grancho M., Beauville L.,
RA Glowaczower C., Lemery D., Ferrara M., Bignon Y.-J.;
RT "The 1396del A mutation and a missense mutation or a rare polymorphism
RT of the WRN gene detected in a French Werner family with a severe
RT phenotype and a case of an unusual vulvar cancer.";
RL Hum. Mutat. 11:413-414(1998).
RN [10]
RP VARIANTS ALA-324 AND ARG-1367.
RX MEDLINE=99167244; PubMed=10069711;

```


RA Castro E., Ogburn C.E., Hunt K.E., Tilvis R., Louhija J., Penttinen R., Erkkola R., Panduro A., Riestra R., Piusens C., Deeb S.S., Wang L., Edland S.D., Martin G.M., Oshima J., RT "Polymorphisms at the Werner locus: I. Newly identified polymorphisms, ethnic variability of 1367Cys/Arg, and its stability in a population of Finnish centenarians.", Am. J. Med. Genet. 82:399-403(1999).

CC -!- FUNCTION: Essential for the formation of DNA replication focal centers; stably associates with foci elements generating binding sites for RP-A. Exhibits a magnesium-dependent ATP-dependent DNA-helicase activity. May be involved in the control of genomic stability (by similarity).

CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.

CC -!- DISEASE: Defects in WRN are the cause of Werner syndrome (WS) [MIM:277700]. WS is a rare autosomal recessive progeroid syndrome characterized by the premature onset of multiple age-related disorders, including atherosclerosis, cancer, non-insulin-dependent diabetes mellitus, ocular cataracts and osteoporosis. The major cause of death, at a median age of 47, is myocardial infarction.

CC -!- SIMILARITY: Belongs to the helicase family. RecQ subfamily.

CC -!- SIMILARITY: Contains 1 HRDC domain.

CC -!- DATABASE: NAME=WRN; NOTE=WRN mutation db (Warner disease); WWW="http://www.pathology.washington.edu/werner/ws wrn.html".

CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.; WWW="http://www.infobiogen.fr/services/chronocancer/Genes/WRNID284.html".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; L76937; AAC41981.1; -.

DR EMBL; AF091214; AACG3361.1; -.

DR EMBL; AF181897; AAF06162.1; -.

DR EMBL; AF181896; AAF06162.1; JOINED.

DR Genew; HGNC:12791; WRN.

DR MIM; 604611; -.

DR MIM; 277700; -.

DR GO; GO:0005634; C:nucleus; TAS.

DR GO; GO:0008408; F:3'-5' exonuclease activity; TAS.

DR GO; GO:0003678; F:DNA helicase activity; TAS.

DR GO; GO:0007568; P:aging; TAS.

DR InterPro; IPR002562; 3_5_exonuclease.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR002121; HRDC.

DR InterPro; IPR004589; RecQ.

DR Pfam; PF01612; 3_5_exonuclease; 1.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase_C; 1.

DR Pfam; PF00570; HRDC; 1.

DR SMART; SM00474; 35EXOC; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELICC; 1.

DR SMART; SM00341; HRDC; 1.

DR TIGRFAMs; TIGR00614; RecQ_fam; 1.

DR PROSITE; PS00690; DEAD ATP HELICASE; FALSE NEG.

KW HydroLase; Helicase; ATP-binding; DNA-binding; Nuclear protein; Polymorphism; Repeat.

FT DOMAIN 424 477 2 X 27 AA TANDem REPEATS OF H-L-S-P-N-D-N-E-N-D-T-S-Y-V-I-E-S-D-E-D-L-E-M-E-M-L-K.

FT REPEAT 424 450 1.

FT REPEAT 451 477 2.

FT DOMAIN 507 510 POLY-GLU.

FT DOMAIN 1150 1229 HRDC.

FT NP_BIND 571 578 ATP (BY SIMILARITY).

FT SITE 668 671 DEAD BOX.

FT VARIANT 324 324 T -> A (in dbSNP:1800390).

FT FTID=VAR_006904. M -> I (in dbSNP:1800391).

FT FTID=VAR_006905. R -> C (in dbSNP:3087425).

FT FTID=VAR_014913. F -> L (in dbSNP:2725362).

FT FTID=VAR_007903. S -> L (in dbSNP:3087414).

FT FTID=VAR_014914. C -> R [polymorphism associated with a higher risk of myocardial infarction; dbSNP:1346044].

FT FTID=VAR_006906. S Q SEQUENCE 1432 AA; 162494 MW; DF02C0059F7B62EB CRC64;

Query Match 72.1%; Score 31; DB 1; Length 1432;

Best Local Similarity 62.5%; Pred. No. 2.6e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8

Db 16 PEWMNVQN 23

RESULT 40

KITH HSV11 STANDARD; PRT; 228 AA.

ID KITH HSV11 STANDARD; PRT; 228 AA.

AC F28855;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Thymidine kinase (EC 2.7.1.21).

GN TK OR 5

OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Ictalurid Herpes-like viruses.

OX NCBI_TaxID=10401;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Auburn 1;

RX MEDLINE=92087490; PubMed=1727613;

RA Davison A.J.;

RT "Channel catfish virus: a new type of herpesvirus.";

RL Virology 186:9-14(1992).

RP [2]

RP DISCUSSION OF SEQUENCE.

RA MEDLINE=92013982; PubMed=1919533;

RX Harrison P.T., Thompson R., Davison A.J.;

RT "Evolution of herpesvirus thymidine kinases from cellular deoxycytidine kinase.";

RL J. Gen. Virol. 72:2583-2586(1991).

CC -!- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-phosphate.

CC -!- SIMILARITY: Belongs to the DCK/DCK family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M75136; AAA88186.1; -.

DR EMBL; M75136; AAA88108.1; -.

DR PIR; JQ1336; KIBEIC.

DR InterPro; IPR002624; DCK.

DR Pfam; PF01712; DCK; 1.

DR Transferase; Kinase; DNA synthesis; ATP-binding.

FT NP_BIND 23 30 ATP (PROBABLE).

SQ SEQUENCE 228 AA; 25642 MW; 980C268797872803 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 228;

```
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYNVNQ 7
Db 85 SYNVNQ 90

RESULT 41
SFSA SYNEL STANDARD; PRT; 241 AA.
ID SFSA SYNEL STANDARD; PRT; 241 AA.
AC Q8DI93;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sugar fermentation stimulation protein homolog.
SFSA OR TLR1697.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpko S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RL Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -1- SIMILARITY: Belongs to the sfsa family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF005374; BAC09249.1; -.
CC HMAP; MF 00095; -.
CC InterPro; IPR005224; Sfsa.
CC Pfam; PF03749; Sfsa; 1.
CC TIGRFAMs; TIGR00230; sfSA; 1.
CC Complete proteome.
KW SEQUENCE 241 AA; 27180 MW; AE44FBEDE9B60F25 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 241;
Best Local Similarity 62.5%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 136 PAVVEVKN 143

RESULT 42
DAPD BUCAI STANDARD; PRT; 274 AA.
ID DAPD BUCAI STANDARD; PRT; 274 AA.
AC P57323;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
DE (EC 2.3.1.117) (Tetrahydropicolinate N-succinyltransferase)
DE (THP succinyltransferase) (Tetrahydropicolinate succinylase).
GN DAPD OR BU229.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
```

```
NCBI_TaxID=118099;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RL Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + (R)-2,3,4,5-tetrahydropyridine-
CC 2,6-dicarboxylate + H(2)O = CoA + (R)-2-(succinylamino)-6-
CC oxoheptanedioate.
CC -1- PATHWAY: Biosynthesis of diaminopimelate and lysine from aspartate
CC semialdehyde; fourth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001118; BAB12944.1; -.
CC HSSP; P56220; 3TDT.
CC InterPro; IPR001451; Hexapep_transf.
CC Pfam; PF00132; hexapep; 5.
CC TIGRFAMs; TIGR00965; ddpd; 1.
CC PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
CC Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis; Complete proteome.
SQ SEQUENCE 274 AA; 31184 MW; E5E923BE9584A936 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 274;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 125 PSYINI 130

RESULT 43
FAC2 DROME STANDARD; PRT; 302 AA.
ID FAC2 DROME STANDARD; PRT; 302 AA.
AC Q9UIH8; Q8SZZ3; Q9VRM4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CAAK prenyl protease 2 (EC 3.4.22.-) (Prenyl protein-specific
DE endoprotease 2) (Farnesylated-proteins converting enzyme 2) (FACE-2)
DE (Severas protein).
GN SRAS OR CG4852.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Weinzierl-Hinun A., Toeroek I., Kiss I., Farkas R., Mechler B.M.;
RT "The severas gene of Drosophila encodes a CAAK-protease and acts as a
RL tumour suppressor.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Fandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimios I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RA [3]
RA REVISIONS
RA MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RA "Annotation of the Drosophila melanogaster euchromatic genome: a
RA systematic review."
RA Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RA [4]
RA SEQUENCE FROM N.A.
RA STRAIN=Berkeley; TISSUE=Embryo;
RA MEDLINE=22426066; PubMed=12537569;
RA George R.A., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA Rubin G.M., Celisner S.E.;
RA "A Drosophila full-length cDNA resource."
RA Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RA CC -!- FUNCTION: Proteolytically removes the C-terminal three residues of
RA CC farnesylated and geranylated proteins (By similarity).
RA CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
RA CC reticulum (By similarity).
RA CC -!- SIMILARITY: Belongs to peptidase family U48.
RA CC
RA CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RA CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA CC the European Bioinformatics Institute. There are no restrictions on its
RA CC use by non-profit institutions as long as its content is in no way
RA CC modified and this statement is not removed. Usage by and for commercial
RA CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA CC or send an email to license@isb-sib.ch).
RA CC

DR EMBL; AJ252068; CAB64383.1; ALT INIT.
DR EMBL; AE003565; AAF50770.3; -
DR EMBL; AY069692; AAL39837.1; -
DR MEROPS; U48.UFW; -
DR FlyBase; FBgn0029121; Stras.
DR InterPro; IPR003675; Abi.
DR Pfam; PF02517; Abi; 1.
KW Hydrolase; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 27 47
FT TRANSMEM 64 84
FT TRANSMEM 104 124
FT TRANSMEM 171 191
FT TRANSMEM 212 232
FT TRANSMEM 236 256
FT TRANSMEM 268 288
FT CONFLICT 94 95
FT CONFLICT 122 122
FT CONFLICT 138 138
SQ SEQUENCE 302 AA; 34415 MW; 84A9E949F1993C3 CRC64;
Query Match 69.8%; Score 30; DB 1; Length 302;
Best Local Similarity 62.5%; Pred. NO. 77;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PSYUNVQN 8
| : : : :
Db 121 PIFVNMQN 128
RESULT 44
NUOD DEIRA
ID NUOD DEIRA STANDARD; PRT; 401 AA.
AC Q9RU089;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-quinone oxidoreductase chain D (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain D) (NDH-1, chain D).
GN NUOD OR DR1503.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
RA radiodurans R1."
RA Science 286:1571-1577(1999).
RA CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
RA CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
RA CC Couples the redox reaction to proton translocation (for every two
RA CC electrons transferred, four hydrogen ions are translocated across
RA CC the cytoplasmic membrane), and thus conserves the redox energy in
RA CC a proton gradient (By similarity).
RA CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
RA CC -!- SIMILARITY: Belongs to the complex I 49 kDa subunit family.
RA CC
RA CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RA CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA CC the European Bioinformatics Institute. There are no restrictions on its
RA CC use by non-profit institutions as long as its content is in no way
RA CC modified and this statement is not removed. Usage by and for commercial
RA CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA CC or send an email to license@isb-sib.ch).
RA CC

```
CC -----
DR EMBL; AE001994; AAFL1069.1; -.
DR PIR; A75388; AV5388.
DR TIGR; DR1503; -.
DR InterPro; IPR001135; Oxidored 49kDa.
DR Pfam; PF00346; complex1_49k; 1.
DR PROSITE; PS00535; COMPLEX1_49K; 1.
DR OXIDOREDUCTASE; NAD; Quinone; Complete proteome.
KW OXIDOREDUCTASE; NAD; Quinone; Complete proteome.
SQ SEQUENCE 401 AA; 44737 MW; C3C5ZB74IA2C684F CRC64;

Query Match 69.8%; Score 30; DB 1; Length 401;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
DB 365 PSFVNLQ 371

RESULT 45
NUD2 RHIME STANDARD; PRT; 404 AA.
AC P56908;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-quinone oxidoreductase chain D 2 (EC 1.6.99.5) (NADH
denydrogenase I, chain D 2) (NDH-1, chain D 2).
GN NUO02 OR RA0831 OR SWA1529.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN STRAIN=1021;
RP SEQUENCE FROM N.A.
RC STRAIN=41;
RA Putnoky P., Jady B., Chellapilla K.P., Barta F., Kiss E.;
RT "Rhizobium meliloti carries two sets of nuo genes.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
sulfur (Fe-S) centers, to quinones in the respiratory chain. The
immediate electron acceptor for the enzyme in this species is
believed to be ubiquinone. Couples the redox reaction to proton
translocation (for every two electrons transferred, four hydrogen
ions are translocated across the cytoplasmic membrane), and thus
conserves the redox energy in a proton gradient (By similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- SIMILARITY: Belongs to the complex I 49 kDa subunit family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; A245399; CAB51632.1; -.
DR EMBL; AE007270; AAK65489.1; -.
DR PIR; G95365; G95365.
```

```
DR InterPro; IPR001135; Oxidored 49kDa.
DR Pfam; PF00346; complex1_49k; 1.
DR PROSITE; PS00535; COMPLEX1_49K; FALSE NEG.
KW OXIDOREDUCTASE; NAD; Quinone; Ubiquinone; Plasmid; Complete proteome.
FT CONFUCTION 295 295
SQ SEQUENCE 404 AA; 45701 MW; 08S527A7D0A4F2BE CRC64;

Query Match 69.8%; Score 30; DB 1; Length 404;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
DB 368 PSFVNLQ 374

RESULT 46
MESJ BUCAI STANDARD; PRT; 440 AA.
AC P57211;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative cell cycle protein mesJ homolog.
GN MESJ OR BU110.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]_TaxID=118099;
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- SIMILARITY: BELONGS TO THE UPF0072 (MESJ/YCF62) FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AP001118; BAB12829.1; -.
DR InterPro; IPR000541; UPF0021.
DR Pfam; PF01171; ATP_bind3; 1.
KW Complete proteome.
SQ SEQUENCE 440 AA; 52630 MW; 46547AC01B27782E CRC64;

Query Match 69.8%; Score 30; DB 1; Length 440;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
DB 268 PSYKNIQ 274

RESULT 47
Y800 METJA STANDARD; PRT; 448 AA.
ID YB00 METJA
AC Q58500;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein MJ1100.
GN MJ1100.
OS Methanococcus jannaschii.
```

```

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.L., Fuhrman J.L., Nguyen D.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton K.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073 (1996).
CC -!- SIMILARITY: Belongs to the phosphohexose mutase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U67553; AAB99103.1; -.
DR PIR; C64437; C64437.
DR TIGR; MJ1100; -.
DR InterPro; IPR005841; PG/PMM mutase.
DR InterPro; IPR005844; PG PMM ABAI.
DR InterPro; IPR005845; PG PMM ABAIL.
DR InterPro; IPR005846; PG PMM ABAILI.
DR InterPro; IPR005843; PG PMM C.
DR Pfam; PF00408; PG PMM; 1.
DR Pfam; PF02878; PG PMM I; 1.
DR Pfam; PF02879; PG PMM II; 1.
DR Pfam; PF02880; PG PMM III; 1.
DR PRINTS; PR00509; PGMPMM.
DR PROSITE; PS00710; PG PMM; 1.
KW Hypothetical protein; Phosphorylation; Complete proteome.
FT ACT_SITE 89 89 PHOSPHOSERINE INTERMEDIATE
FT ACT_SITE 89 89 (BY SIMILARITY).
SQ SEQUENCE 448 AA; 50098 MW; FBC7EF17A73DF9B4 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 448;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 357 PSYVNL 363
ID UBP8 YEAST STANDARD; PRT; 471 AA.
AC P50102;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase 8 (EC 3.1.2.15) (Ubiquitin
DE thiolesterase 8) (Ubiquitin-specific processing protease 8)
DE (Deubiquitinating enzyme 8).
GN UBP8 OR YMR223W OR YMR959.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93 (1997).
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -!- SIMILARITY: Belongs to peptidase family C19.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z49939; CAA90194.1; -.
DR PIR; S57591; S57591.
DR GenOnline; 142898; -.
DR MEROPS; C19.UFW; -.
DR SGD; S0004836; UBP8.
DR GO; GO:0000124; C:SAGA complex; IDA.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00443; UCH; 1.
DR Pfam; PF02148; zf-UBP; 1.
DR SMART; SMO0290; Znf_UBP; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 146 146 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
FT ACT_SITE 427 427 BY SIMILARITY.
SQ SEQUENCE 471 AA; 53623 MW; BC632F12FBD0F73C CRC64;

Query Match 69.8%; Score 30; DB 1; Length 471;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8
Db 383 PTYLNKN 390
ID HEMA_IADL1 STANDARD; PRT; 565 AA.
AC P04661;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA
OS Influenza A virus (strain A/Duck/England/1/56).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11354;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220697; PubMed=2024485;
RA Nobusawa E., Aoyama T., Kato H., Suzuki Y., Tateno Y., Nakajima K.;
RT "Comparison of complete amino acid sequences and receptor-binding
RT properties among 13 serotypes of hemagglutinins of influenza A
RT viruses.";
RL Virology 182:475-485 (1991).
RN [2]

```

RP SEQUENCE OF 1-83 FROM N.A.
 RX MEDLINE=82150925; PubMed=6174976;
 RA Air G.M.;
 RA "Sequence relationships among the hemagglutinin genes of 12 subtypes
 of influenza A virus.";
 RT Proc. Natl. Acad. Sci. U.S.A. 78:7639-7643 (1981).
 CC -!- FUNCTION: Hemagglutinin is responsible for attaching the virus to
 cell receptors and for initiating infection.
 CC -!- SUBUNIT: Homotrimer. Each of the monomers is formed by two chains
 (HA1 and HA2) linked by a disulfide bond.
 CC -!- SIMILARITY: Belongs to the influenza viruses hemagglutinin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: D90306; BAA14316.1; -;
 DR EMBL: J02107; BAA43183.1; -;
 DR HSP: P03437; IHTM.
 DR InterPro: IPR008980; Capsid_hemag.
 DR Pfam: PF001364; Hemagglutn.
 DR PRINTS: PR00329; HEMAGGLUTN12.
 DR ProDom: PD000225; Hemagglutn; 1.
 KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 341 HEMAGGLUTININ HA1 CHAIN.
 FT CHAIN 343 565 HEMAGGLUTININ HA2 CHAIN.
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 496 496 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 4 4 T -> I (IN REF. 2).
 SQ SEQUENCE 565 AA; 63097 MW; 81894320E70C908F CRC64;

Query Match 69.8%; Score 30; DB 1; Length 565;
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
 DB 319 PKYVNVKS 326

RESULT 50
 HP11 METAC STANDARD; PRT; 676 AA.
 ID HP11 METAC
 AC Q8TVA9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pyrophosphate-energized proton pump 1 (EC 3.6.1.1) (Pyrophosphate-
 energized inorganic pyrophosphatase 1) (H+-Ppase 1) (Membrane-bound
 proton-translocating pyrophosphatase 1).
 DE HPA1 OR MA3879.
 GN Methanosarcina acetivorans.
 OS Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_taxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=1192238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RL Genome Res. 12:532-542 (2002).
 CC -!- FUNCTION: Generates a proton motive force; it probably catalyzes a
 fully reversible reaction, thus being able to synthesize
 CC pyrophosphate when the proton motive force is sufficient (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
 CC -!- COFACTOR: Magnesium and potassium (By similarity).
 CC -!- SUBUNIT: Homodimer (Potential).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: Belongs to the H(+)-translocating pyrophosphatase
 (TC 3.A.10) family. Subfamily 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AB011099; AAM07230.1; ALT_INIT.
 DR HAMAP: MF 01129; -; 1.
 DR InterPro: IPR004131; H_Ppase.
 DR Pfam: PF03030; H_Ppase; 1.
 DR TIGRFAMs: TIGR01104; V_Ppase; 1.
 KW Hydrogen ion transport; Hydrolase; Magnesium; Potassium;
 Transmembrane; Complete proteome.
 FT TRANSMEM 4 23 Potential.
 FT TRANSMEM 57 79 Potential.
 FT TRANSMEM 84 106 Potential.
 FT TRANSMEM 127 149 Potential.
 FT TRANSMEM 159 181 Potential.
 FT TRANSMEM 237 259 Potential.
 FT TRANSMEM 264 286 Potential.
 FT TRANSMEM 299 321 Potential.
 FT TRANSMEM 326 348 Potential.
 FT TRANSMEM 389 411 Potential.
 FT TRANSMEM 472 494 Potential.
 FT TRANSMEM 501 523 Potential.
 FT TRANSMEM 568 587 Potential.
 FT TRANSMEM 594 616 Potential.
 FT TRANSMEM 653 675 Potential.
 FT SITE 463 463 DETERMINANT OF POTASSIUM DEPENDENCE (BY
 SIMILARITY).
 SQ SEQUENCE 676 AA; 69276 MW; 5C7144B63F1F8E26 CRC64;

Query Match 69.8%; Score 30; DB 1; Length 676;
 Best Local Similarity 85.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SYVNVQN 8
 DB 22 SYKVVQN 28

Search completed: May 24, 2004, 17:31:31
 Job time : 24 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:28:52 ; Search time 40 Seconds
 (without alignments)
 63.104 Million cell updates/sec

Title: US-09-977-349-5
 Perfect score: 43
 Sequence: 1 PSYVNVQN 8

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
 Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 150 summaries

Database :

SPTREMBL.25.*
 1: sp_archaea.*
 2: sp_bacteria.*
 3: sp_fungi.*
 4: sp_human.*
 5: sp_invertebrate.*
 6: sp_mammal.*
 7: sp_mhc.*
 8: sp_organelle.*
 9: sp_phase.*
 10: sp_plant.*
 11: sp_rodent.*
 12: sp_virus.*
 13: sp_vertibrate.*
 14: sp_unclassified.*
 15: sp_rvirus.*
 16: sp_bacteriap.*
 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	369	4	Q8N4K5 homo sapien
2	43	100.0	470	13	Q8AY68 xenopus lae
3	43	100.0	474	4	Q96CL1
4	42	97.7	469	11	Q8BFY3
5	37	86.0	278	12	Q8JSS6
6	37	86.0	1338	5	Q86A39
7	35	81.4	123	10	Q40821
8	35	81.4	536	5	Q95PJ5
9	34	79.1	195	2	Q8GA81
10	34	79.1	392	3	Q94431
11	34	79.1	6088	5	Q8IEM1
12	33	76.7	105	15	Q9Q6K6
13	33	76.7	105	15	Q9Q6K2
14	33	76.7	105	15	Q9Q6K0
15	33	76.7	105	15	Q9Q6M7
16	33	76.7	105	15	Q9Q6M2

17	33	76.7	105	15	Q9Q6L8	Q9Q6L8 human t-lym
18	33	76.7	105	15	Q9Q6J1	Q9Q6J1 human t-lym
19	33	76.7	105	15	Q9Q6L7	Q9Q6L7 human t-lym
20	33	76.7	105	15	Q9Q6K4	Q9Q6K4 human t-lym
21	33	76.7	105	15	Q9Q6L5	Q9Q6L5 human t-lym
22	33	76.7	105	15	Q9Q6M6	Q9Q6M6 human t-lym
23	33	76.7	105	15	Q9Q6J6	Q9Q6J6 human t-lym
24	33	76.7	105	15	Q9Q6L4	Q9Q6L4 human t-lym
25	33	76.7	105	15	Q9Q6J2	Q9Q6J2 human t-lym
26	33	76.7	105	15	Q9Q6K5	Q9Q6K5 human t-lym
27	33	76.7	105	15	Q9Q6N0	Q9Q6N0 human t-lym
28	33	76.7	105	15	Q9Q6J8	Q9Q6J8 human t-lym
29	33	76.7	105	15	Q9Q6M0	Q9Q6M0 human t-lym
30	33	76.7	105	15	Q9Q6L9	Q9Q6L9 human t-lym
31	33	76.7	105	15	Q9Q6K1	Q9Q6K1 human t-lym
32	33	76.7	105	15	Q9Q6K8	Q9Q6K8 human t-lym
33	33	76.7	105	15	Q9Q6J4	Q9Q6J4 human t-lym
34	33	76.7	105	15	Q9Q6K7	Q9Q6K7 human t-lym
35	33	76.7	105	15	Q9Q6K9	Q9Q6K9 human t-lym
36	33	76.7	105	15	Q9Q6J3	Q9Q6J3 human t-lym
37	33	76.7	105	15	Q9Q6L1	Q9Q6L1 human t-lym
38	33	76.7	105	15	Q9Q6M4	Q9Q6M4 human t-lym
39	33	76.7	105	15	Q9Q6M5	Q9Q6M5 human t-lym
40	33	76.7	105	15	Q9Q6J9	Q9Q6J9 human t-lym
41	33	76.7	105	15	Q9Q6L3	Q9Q6L3 human t-lym
42	33	76.7	105	15	Q9Q6J7	Q9Q6J7 human t-lym
43	33	76.7	105	15	Q9Q6M1	Q9Q6M1 human t-lym
44	33	76.7	105	15	Q9Q6N1	Q9Q6N1 human t-lym
45	33	76.7	105	15	Q9Q6J5	Q9Q6J5 human t-lym
46	33	76.7	105	15	Q9Q6L0	Q9Q6L0 human t-lym
47	33	76.7	105	15	Q9Q6L6	Q9Q6L6 human t-lym
48	33	76.7	105	15	Q9Q6M8	Q9Q6M8 human t-lym
49	33	76.7	105	15	Q9Q6K3	Q9Q6K3 human t-lym
50	33	76.7	105	15	Q9Q6M3	Q9Q6M3 human t-lym
51	33	76.7	105	15	Q9Q6L2	Q9Q6L2 human t-lym
52	33	76.7	105	15	Q9Q6N2	Q9Q6N2 human t-lym
53	33	76.7	105	15	Q9Q6N3	Q9Q6N3 human t-lym
54	33	76.7	125	4	Q96A04	Q96A04 homo sapien
55	33	76.7	297	5	Q8I4T1	Q8I4T1 plasmodium
56	33	76.7	322	4	Q86V15	Q86V15 homo sapien
57	33	76.7	409	16	Q8FRQ9	Q8FRQ9 corynebacte
58	33	76.7	527	17	Q29232	Q29232 archaeoglob
59	33	76.7	534	5	Q27465	Q27465 caenorhabdi
60	33	76.7	577	16	Q99RK4	Q99RK4 staphylococ
61	33	76.7	577	16	Q8NVQ3	Q8NVQ3 staphylococ
62	33	76.7	588	16	Q67183	Q67183 aquifex aeo
63	33	76.7	731	15	Q82324	Q82324 human t-lym
64	33	76.7	732	5	Q24557	Q24557 drosophila
65	33	76.7	732	5	Q24558	Q24558 drosophila
66	33	76.7	732	5	Q960X5	Q960X5 drosophila
67	33	76.7	776	16	Q98PT2	Q98PT2 mycoplasma
68	33	76.7	896	15	Q56622	Q56622 human t-lym
69	33	76.7	896	15	Q82232	Q82232 human t-lym
70	33	76.7	910	10	Q9C905	Q9C905 arabidopsis
71	33	76.7	965	10	Q9C6A8	Q9C6A8 arabidopsis
72	33	76.7	990	17	Q97VM0	Q97VM0 sulfolobus
73	33	76.7	1044	2	Q9F4D6	Q9F4D6 bifidobacte
74	33	76.7	1167	17	Q8TUL9	Q8TUL9 methanosarc
75	33	76.7	1462	15	Q56228	Q56228 human t-lym
76	33	76.7	4345	5	Q9VLA0	Q9VLA0 drosophila
77	32	74.4	99	16	Q8EJS2	Q8EJS2 shewanella
78	32	74.4	145	16	Q89PR6	Q89PR6 bradyrhizob
79	32	74.4	202	12	Q80MP4	Q80MP4 indian citr
80	32	74.4	215	5	Q18661	Q18661 caenorhabdi
81	32	74.4	222	12	Q9QEE6	Q9QEE6 indian citr
82	32	74.4	450	16	Q8XIG7	Q8XIG7 clostridium
83	32	74.4	469	8	Q85YM9	Q85YM9 euphorbia e
84	32	74.4	477	16	Q892C8	Q892C8 clostridium
85	32	74.4	538	4	Q7Z7G3	Q7Z7G3 homo sapien
86	32	74.4	542	16	Q81TD6	Q81TD6 bacillus an
87	32	74.4	548	2	Q9RDZ3	Q9RDZ3 legionella
88	32	74.4	565	10	Q84J37	Q84J37 arabidopsis
89	32	74.4	576	5	Q8ID23	Q8ID23 plasmodium

90 Q9npb5 homo sapien
 91 Q9nux6 ciona savig
 92 O02427 ciona intes
 93 Q7zys8 brachydanio
 94 Q9uv19 blumeria gr
 95 Q81g77 bacillus ce
 96 Q9hge4 arthroderna
 97 Q8tgd4 aspergillus
 98 Q9p442 coccidioid
 99 Q94165 emericella
 100 Q8asn6 bacteroides
 101 Q875j0 candida alb
 102 Q8ken4 chlorobium
 103 Q81eq9 bacillus ce
 104 Q7yvr5 cryptospori
 105 Q40805 physalis cr
 106 Q96xw0 sulfolobus
 107 Q94g36 physalis lo
 108 Q8rws0 arabidopsis
 109 Q9ye10 aeropyrum p
 110 Q9uws5 sulfolobus
 111 Q8d4c6 vibrio vuln
 112 Q8xi25 clostridium
 113 Q8dids vibrio vuln
 114 Q87rq7 vibrio para
 115 Q9kl52 neisseria m
 116 Q9jss7 neisseria m
 117 Q82a90 nitrosomona
 118 Q7w8y2 bordetella
 119 Q9cmz2 pascuella
 120 Q8zrp4 salmonella
 121 Q8zh69 yersinia pe
 122 Q8x8y7 escherichia
 123 Q8fl11 escherichia
 124 Q829a8 salmonella
 125 Q7vnc4 haemophilus
 126 Q8xzk2 ralstonia s
 127 Q7vres candidatus
 128 Q8czy9 yersinia pe
 129 Q82t62 nitrosomona
 130 Q87vm3 pseudomonas
 131 Q8icq1 plasmidium
 132 Q897a2 clostridium
 133 Q62443 methanobact
 134 Q65263 arabidopsis
 135 Q9fix8 arabidopsis
 136 Q93j06 human herpe
 137 Q9f773 campylobact
 138 Q8ibq7 plasmidium
 139 Q9u6m5 globodera t
 140 Q9pnc5 campylobact
 141 Q16028 globodera r
 142 Q8mlg0 drosophila
 143 Q7v8q0 prochloroco
 144 Q8on77 populus tri
 145 Q8yzc2 anabaena sp
 146 Q8yz76 anabaena sp
 147 Q39717 nicotiana t
 148 Q8xkf5 clostridium
 149 Q93695 sulfolobus
 150 Q9c4v5 sulfolobus

ALIGNMENTS

RESULT 1
 Q8N4K5 PRELIMINARY; PRT; 369 AA.
 ID Q8N4K5
 AC Q8N4K5; (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RC TISSUE=Brain;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC033925; AAH3925.1; -.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR006019; PID domain.
 DR InterPro; IPR006020; PTB_PID.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00640; PID; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00629; SHCPIDOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00462; PTB; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS01179; PID; 1.
 DR PROSITE; PS00001; SH2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 369 AA; 40414 MW; 46F34449B556DDDD0 CRC64;
 Query Match 100.0%; Score 43; DB 4; Length 369;
 Best Local Similarity 100.0%; Pred.No.1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PSYVNVQN 8
 Db 211 PSYVNVQN 218
 RESULT 2
 Q8AY68 PRELIMINARY; PRT; 470 AA.
 ID Q8AY68
 AC Q8AY68;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Src homology collagen.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chesnel F., Heligon C., Boujard D.;
 RT "Molecular cloning and characterization of the adaptor molecule Shc
 from Xenopus laevis oocytes: possible implication in mediating
 insulin-induced meiotic resumption."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY027793; AAK14789.1; -.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR006019; PID domain.
 DR InterPro; IPR006020; PTB_PID.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00640; PID; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00629; SHCPIDOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00462; PTB; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS01179; PID; 1.
 DR PROSITE; PS00001; SH2; 1.
 KW Collagen.
 SQ SEQUENCE 470 AA; 52367 MW; FF8EBBBA9FE446B1 CRC64;


```

Query Match      100.0%; Score 43; DB 13; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYNNVQN 8
DB      323 PSYNNVQN 330

RESULT 3
Q96CL1 ID Q96CL1 PRELIMINARY; PRT; 474 AA.
AC Q96CL1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014158; AAH14158.1; -.
DR GO; GO:0007243; P:intracellular signaling cascade; IEA.
DR InterPro; IPR006019; PID domain.
DR InterPro; IPR006020; PTB PID.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00640; PID; 1.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00462; PTB; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS01179; PID; 1.
DR PROSITE; PS50001; SH2; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51681 MW; 4381D98CBA87DB37 CRC64;

Query Match      100.0%; Score 43; DB 4; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYNNVQN 8
DB      316 PSYNNVQN 323

RESULT 4
Q8BFY3 ID Q8BFY3 PRELIMINARY; PRT; 469 AA.
AC Q8BFY3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Src homology 2 domain-containing transforming protein C1.
GN SHC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;

```

```

RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; BC036172; AAH36172.1; -.
DR EMBL; AK049357; BAC33706.1; -.
DR FIR; A55484; A55484.
DR MGD; MGI:98296; Shc1.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR006019; PID domain.
DR InterPro; IPR006020; PTB PID.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00640; PID; 1.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00462; PTB; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS01179; PID; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 469 AA; 51409 MW; DF33594E80107A3E CRC64;

Query Match      97.7%; Score 42; DB 11; Length 469;
Best Local Similarity 87.5%; Pred. No. 3.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYNNVQN 8
DB      311 PSYNNVQN 318

RESULT 5
Q8JS56 ID Q8JS56 PRELIMINARY; PRT; 278 AA.
AC Q8JS56;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein kinase 1.
GN PHOGV003.
OS Phthorimaea operculella granulovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=192584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21011882; PubMed=11129630;
RA Taba A., Nour-El-Din A., Croizier L., Ferber M.L., Croizier G.;
RT "Comparative analysis of the granulin regions of the Phthorimaea
  operculella and Spodoptera litoralis granuloviruses.";
RL Virus Genes 21:147-155 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Croizier L., Taba A., Croizier G., Lopez Ferber M.;
RT "The complete sequence of the potato tuber moth, Phthorimaea
  operculella, granulovirus.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF499596; AAM70201.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS50011; PROTEIN KINASE DM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
ATP-binding; Kinase; transferase.
KW

```

```

SQ SEQUENCE 278 AA; 32481 MW; 45AB9FOCE0AF5FE CRC64;
Query Match 86.0%; Score 37; DB 12; Length 278;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
Db 71 PHYVNIQN 78

RESULT 6
Q86A39 PRELIMINARY; PRT; 1338 AA.
AC Q86A39;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). protein kinase
DE Yaka.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunga B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116988; AAC052003.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR000313; PWWP_domain.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00293; PWWP; 1.
KW Kinase.
SQ SEQUENCE 1338 AA; 150907 MW; 623F29A745907CFB CRC64;

Query Match 86.0%; Score 37; DB 5; Length 1338;
Best Local Similarity 75.0%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
Db 900 PHYVNIQN 907

RESULT 7
Q40821 PRELIMINARY; PRT; 123 AA.
AC Q40821;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S-RNase (Fragment).
GN S.
OS Physalis crassifolia (Ground cherry).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Physalis.
OX NCBI_TaxID=49773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96224705; PubMed=8666546;
RA Richman A.D., Uyenoyma M.K., Kohn J.R.;
RT "S-allele diversity in a natural population of Physalis crassifolia

RT (Solanaceae) (ground cherry) assessed by RT-PCR.";
RL Heredit 76:497-505(1996).
DR EMBL; L46657; AAB37222.1; -.
DR GO; GO:0004521; F:endoribonuclease activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR001568; RNase_T2.
DR Pfam; PF00445; ribonuclease_T2; 1.
DR PROSITE; PS00531; RNASE_T2_2; 1.
FT NON_TER 1
FT TER 123
SQ SEQUENCE 123 AA; 14260 MW; 0CD6ECC005E1CA06 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 123;
Best Local Similarity 62.5%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
Db 13 PHYVNIQN 20

RESULT 8
Q95PJ5 PRELIMINARY; PRT; 536 AA.
AC Q95PJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Calnexin.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RA Ecke M., Gerisch G.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF073837; AAKS8500.1; -.
DR GO; GO:0005514; F:calcium ion storage activity; IEA.
DR InterPro; IPR001580; Calreticulin.
DR InterPro; IPR009033; Calret calnex P.
DR InterPro; IPR008985; ConA like lec_gl.
DR Pfam; PF00262; calreticulin; 1-__gl.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
SQ SEQUENCE 536 AA; 60479 MW; 850F4F6F830F8595 CRC64;

Query Match 81.4%; Score 35; DB 5; Length 536;
Best Local Similarity 75.0%; Pred. No. 96;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|:|:|:|
Db 350 PSYFNVEN 357

RESULT 9
Q8GA81 PRELIMINARY; PRT; 195 AA.
AC Q8GA81;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=536;
RX MEDLINE=22267134; PubMed=12379716;

```

RA Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA Gottschalk G., Hacker J.;
RT "Genetic structure and distribution of four pathogenicity islands (PAI
RT I536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL Infect. Immun. 70:6365-6372(2002).
DR EMBL; AJ488511; CAD33715.1; -.
KW Hypothetical protein.
SQ SEQUENCE 195 AA; 22396 MW; 9377BCE8AB0C77E5 CRC64;

Query Match 79.1%; Score 34; DB 2; Length 195;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||:|:
Db 179 PSFINIEN 186

RESULT 10
O94431 PRELIMINARY; PRT; 392 AA.
AC O94431;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN SPBC660.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Lyne M., Rajandream M.A., Barrell B.G., Rieger M;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034563; CAA22532.1; -.
DR PIR; T40624; T40624.
DR GeneDB SPombe; SPBC660.12C; -.
KW Hypothetical protein.
SQ SEQUENCE 392 AA; 44535 MW; F4594C9A9DF42913 CRC64;

Query Match 79.1%; Score 34; DB 3; Length 392;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||:|:
Db 19 PDYVNVNN 26

RESULT 11
Q81EN1 PRELIMINARY; PRT; 6088 AA.
AC Q81EN1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN MALL3P1.39.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52218.1; -.
KW Hypothetical protein.
SQ SEQUENCE 6088 AA; 730357 MW; 925F4269FEE27820 CRC64;

Query Match 79.1%; Score 34; DB 5; Length 6088;
Best Local Similarity 75.0%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||:|:
Db 4872 PSYVNTNN 4879

RESULT 12
Q9Q6K6 PRELIMINARY; PRT; 105 AA.
AC Q9Q6K6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN080;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil".
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197312; AAF17534.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||:|:
Db 29 PSYINTDN 36

RESULT 13
Q9Q6K2 PRELIMINARY; PRT; 105 AA.
AC Q9Q6K2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN087;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil".
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197316; AAF17538.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.

```

DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 15; Length 105;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 14
QY Q906K0 PRELIMINARY; PRT; 105 AA.
AC Q906K0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN091;
RA Alcantara L.C.J.; Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197318; AAF17540.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
DR NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 15; Length 105;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 15
QY Q906M7 PRELIMINARY; PRT; 105 AA.
AC Q906M7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN026;
RA Alcantara L.C.J.; Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from

```

```

RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197291; AAF17513.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
DR NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 15; Length 105;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 16
QY Q906M2 PRELIMINARY; PRT; 105 AA.
AC Q906M2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN039;
RA Alcantara L.C.J.; Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197296; AAF17518.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
DR NON_TER 105 105
SQ SEQUENCE 105 AA; 11568 MW; 929C4C3A3F7D46C0 CRC64;

Query Match
Best Local Similarity 76.7%; Score 33; DB 15; Length 105;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 17
QY Q906L8 PRELIMINARY; PRT; 105 AA.
AC Q906L8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=POL-FNN051;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197300; AAF17522.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 DR NON_TER 1 105
 FT NON_TER 1 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;
 Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
 Db 29 PSYINTDN 36

RESULT 18
 Q906J1 PRELIMINARY; PRT; 105 AA.
 AC Q906J1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OC NCBI_TaxID=11908;
 RN [1]
 RC STRAIN=GO110;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Central West from Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197327; AAF17549.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 DR NON_TER 1 105
 FT NON_TER 1 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;
 Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
 Db 29 PSYINTDN 36

RESULT 19
 Q906L7 PRELIMINARY; PRT; 105 AA.
 AC Q906L7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.

OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OC NCBI_TaxID=11908;
 RN [1]
 RC STRAIN=POL-FNN053;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197301; AAF17523.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 DR NON_TER 1 105
 FT NON_TER 1 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
 Db 29 PSYINTDN 36

RESULT 20
 Q906K4 PRELIMINARY; PRT; 105 AA.
 AC Q906K4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OC NCBI_TaxID=11908;
 RN [1]
 RC STRAIN=POL-FNN082;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197314; AAF17536.1; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 DR NON_TER 1 105
 FT NON_TER 1 105
 SQ SEQUENCE 105 AA; 11564 MW; 882DE7208BC3F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
 Db 29 PSYINTDN 36

RESULT 21
 Q906L5 PRELIMINARY; PRT; 105 AA.
 ID Q906L5
 AC Q906L5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OX NCBI_TaxID=11908;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=POL-FNN057;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197303; AAF17525.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 FT NON_TER 1
 FT NON_TER 105
 FT NON_TER 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PSYVNVQN 8
 Db 29 PSYINTDN 36

RESULT 22
 Q9Q6M6 PRELIMINARY; PRT; 105 AA.
 AC Q9Q6M6;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OX NCBI_TaxID=11908;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=POL-FNN028;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197292; AAF17514.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 FT NON_TER 1
 FT NON_TER 105
 FT NON_TER 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PSYVNVQN 8
 Db 29 PSYINTDN 36

RESULT 23
 Q9Q6J6 PRELIMINARY; PRT; 105 AA.
 AC Q9Q6J6;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OX NCBI_TaxID=11908;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TP97/35;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197322; AAF17544.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 FT NON_TER 1
 FT NON_TER 105
 FT NON_TER 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PSYVNVQN 8
 Db 29 PSYINTDN 36

RESULT 24
 Q9Q6L4 PRELIMINARY; PRT; 105 AA.
 AC Q9Q6L4;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polymerase (Fragment).
 GN POL.
 OS Human T-lymphotropic virus 1.
 OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
 OX NCBI_TaxID=11908;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=POL-FNN060;
 RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
 RA Galvao-Castro B.;
 RT "Identification and Phylogenetic Characterization of Human T-
 RT lymphotropic virus type I in the Northeast and Southeast from
 RT Brazil.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF197304; AAF17526.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 FT NON_TER 1
 FT NON_TER 105
 FT NON_TER 105
 SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
 Best Local Similarity 62.5%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PSYVNVQN 8
 Db 29 PSYINTDN 36

```

QY      1 PSYVNVQN 8
Db      29 PSYINTDN 36

RESULT 25
Q06J2 ID Q06J2 PRELIMINARY; PRT; 105 AA.
AC Q06J2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TP98/70;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197326; AAF17548.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      29 PSYINTDN 36

RESULT 26
Q06K5 ID Q06K5 PRELIMINARY; PRT; 105 AA.
AC Q06K5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN081;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197313; AAF17535.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

```

```

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      29 PSYINTDN 36

RESULT 27
Q06N0 ID Q06N0 PRELIMINARY; PRT; 105 AA.
AC Q06N0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN009;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197288; AAF17510.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      29 PSYINTDN 36

RESULT 28
Q06J8 ID Q06J8 PRELIMINARY; PRT; 105 AA.
AC Q06J8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RNT97/38;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197320; AAF17542.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.

```

```
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11552 MW; 882DF371CBD7F6C1 CRC64;

Query Match          76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 29
QYQ6M0 PRELIMINARY; PRT; 105 AA.
AC QYQ6M0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN043;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197298; AAF17520.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match          76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 30
QYQ6L9 PRELIMINARY; PRT; 105 AA.
AC QYQ6L9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN047;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
```

```
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197299; AAF17521.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match          76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 31
QYQ6K1 PRELIMINARY; PRT; 105 AA.
AC QYQ6K1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN089;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197317; AAF17539.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match          76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
Db 29 PSYINTDN 36

RESULT 32
QYQ6K8 PRELIMINARY; PRT; 105 AA.
AC QYQ6K8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
```



```

RC STRAIN=POL-FNN073;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197310; AAF17532.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db |||:|
29 PSYINTDN 36

RESULT 33
Q9Q6J4 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TP98/55;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197324; AAF17546.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db |||:|
29 PSYINTDN 36

RESULT 34
Q9Q6K7 PRELIMINARY; PRT; 105 AA.
AC Q9Q6K7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.

```

```

OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN078;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197311; AAF17533.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db |||:|
29 PSYINTDN 36

RESULT 35
Q9Q6K9 PRELIMINARY; PRT; 105 AA.
AC Q9Q6K9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN072;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197309; AAF17531.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db |||:|
29 PSYINTDN 36

RESULT 36
Q9Q6J3 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J3;

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TP98/66;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197325; AAF17547.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
FT SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;
Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYUNVQN 8
DB 29 PSYINTDN 36
RESULT 37
QYQ6L1
ID QYQ6L1 PRELIMINARY; PRT; 105 AA.
AC QYQ6L1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN068;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197307; AAF17529.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
FT SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;
Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYUNVQN 8
DB 29 PSYINTDN 36

```

```

RESULT 38
QYQ6M4
ID QYQ6M4 PRELIMINARY; PRT; 105 AA.
AC QYQ6M4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN032;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197294; AAF17516.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
FT SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;
Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 PSYUNVQN 8
DB 29 PSYINTDN 36
RESULT 39
QYQ6M5
ID QYQ6M5 PRELIMINARY; PRT; 105 AA.
AC QYQ6M5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN029;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197293; AAF17515.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1
FT NON_TER 105
FT SEQUENCE 105 AA; 11568 MW; 929C4C3A3F7D46C0 CRC64;
Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 40
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6J9 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN094;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197319; AAF17541.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 41
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6L3 PRELIMINARY; PRT; 105 AA.
AC Q9Q6L3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN061;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197305; AAF17527.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 42
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6J7 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=TP97/33;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197321; AAF17543.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 43
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6M1 PRELIMINARY; PRT; 105 AA.
AC Q9Q6M1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN041;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197297; AAF17519.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
```

```
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 42
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6J7 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=TP97/33;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197321; AAF17543.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
DR NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

RESULT 43
QY 1 PSYVNVQV 8
Db 29 PSYINTDN 36

ID Q9Q6M1 PRELIMINARY; PRT; 105 AA.
AC Q9Q6M1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]_TaxID=11908;
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN041;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197297; AAF17519.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
```

```

DR GO: 0006310; P:DNA recombination; IEA.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00665; rve; 1.
FT NON TER 1
FT NON TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8
Db 29 PSYINTDN 36

RESULT 44
Q9Q6N1 ID Q9Q6N1 PRELIMINARY; PRT; 105 AA.
AC Q9Q6N1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GO263;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Central West from Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF197287; AAF17509.1; -.
DR GO: 0006310; P:DNA binding; IEA.
DR GO: 0006310; P:DNA recombination; IEA.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00665; rve; 1.
FT NON TER 1
FT NON TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8
Db 29 PSYINTDN 36

RESULT 45
Q9Q6J5 ID Q9Q6J5 PRELIMINARY; PRT; 105 AA.
AC Q9Q6J5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TP97/36;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from

```

```

RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF197323; AAF17545.1; -.
DR GO: 0003677; F:DNA binding; IEA.
DR GO: 0006310; P:DNA recombination; IEA.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00665; rve; 1.
FT NON TER 1
FT NON TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8
Db 29 PSYINTDN 36

RESULT 46
Q9Q6L0 ID Q9Q6L0 PRELIMINARY; PRT; 105 AA.
AC Q9Q6L0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN071;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF197308; AAF17530.1; -.
DR GO: 0003677; F:DNA binding; IEA.
DR GO: 0006310; P:DNA recombination; IEA.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00665; rve; 1.
FT NON TER 1
FT NON TER 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8
Db 29 PSYINTDN 36

RESULT 47
Q9Q6L6 ID Q9Q6L6 PRELIMINARY; PRT; 105 AA.
AC Q9Q6L6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.

```

```
RC STRAIN=POL-FNN054;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197302; AAF17524.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11572 MW; 882DE7208ED439C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNQV 8
Db 29 PSYINTDN 36

RESULT 48
Q9Q6M8 PRELIMINARY; PRT; 105 AA.
AC Q9Q6M8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN022;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197290; AAF17512.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNQV 8
Db 29 PSYINTDN 36

RESULT 49
Q9Q6K3 PRELIMINARY; PRT; 105 AA.
AC Q9Q6K3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
```

```
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN083;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197315; AAF17537.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNQV 8
Db 29 PSYINTDN 36

RESULT 50
Q9Q6M3 PRELIMINARY; PRT; 105 AA.
AC Q9Q6M3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polymerase (Fragment).
GN POL.
OS Human T-lymphotropic virus 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POL-FNN035;
RA Alcantara L.C.J., Shindo N., Kashima S., Cristina M.C.R., Covas D.T.,
RA Galvao-Castro B.;
RT "Identification and Phylogenetic Characterization of Human T-
RT lymphotropic virus type I in the Northeast and Southeast from
RT Brazil.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF197295; AAF17517.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11582 MW; 882DE7208ED7F6C1 CRC64;

Query Match 76.7%; Score 33; DB 15; Length 105;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNQV 8
Db 29 PSYINTDN 36

Search completed: May 24, 2004, 17:33:04
Job time : 59 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:29:57 ; Search time 23 seconds
(without alignments)
17.957 Million cell updates/sec

Title: US-09-977-349-5
Perfect score: 43
Sequence: 1 PSYVNVQN 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents, AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	12	4	US-09-347-926-18
2	43	100.0	423	1	US-08-353-550-3
3	43	100.0	423	2	US-08-551-687-3
4	43	100.0	473	1	US-08-363-215-2
5	43	100.0	473	3	US-08-807-342B-7
6	43	100.0	473	3	US-08-664-962B-4
7	43	100.0	473	3	US-08-311-743-4
8	43	100.0	474	1	US-08-353-550-8
9	43	100.0	474	2	US-08-551-687-8
10	42	97.7	469	1	US-08-807-342B-6
11	42	97.7	469	3	US-08-363-215-1
12	37	86.0	9	1	US-08-266-514-27
13	37	86.0	9	2	US-08-654-604-27
14	35	81.4	8	1	US-08-408-604A-157
15	35	81.4	9	1	US-08-408-604A-156
16	35	81.4	16	1	US-08-408-604A-154
17	35	81.4	282	4	US-09-328-352-6758
18	34	79.1	251	4	US-09-107-532A-4381
19	34	79.1	273	4	US-09-328-352-5843
20	33	76.7	121	4	US-09-621-976-5040
21	33	76.7	125	4	US-09-621-976-7453
22	31	72.1	71	2	US-08-726-306A-37
23	31	72.1	274	4	US-09-543-681A-6646
24	31	72.1	284	4	US-09-489-039A-12381
25	31	72.1	285	4	US-09-328-352-6357
26	31	72.1	318	4	US-09-252-991A-22236
27	31	72.1	408	4	US-09-107-532A-3913
28	29	67.4	100		Sequence 3913, Ap
29	67.4	97.2	3		Sequence 12381, A
30	67.4	97.2	3		Sequence 6357, Ap
31	67.4	97.2	3		Sequence 12381, A
32	67.4	97.2	3		Sequence 6357, Ap
33	67.4	97.2	3		Sequence 12381, A
34	67.4	97.2	3		Sequence 6357, Ap
35	67.4	97.2	3		Sequence 12381, A
36	67.4	97.2	3		Sequence 6357, Ap
37	67.4	97.2	3		Sequence 12381, A
38	67.4	97.2	3		Sequence 6357, Ap
39	67.4	97.2	3		Sequence 12381, A
40	67.4	97.2	3		Sequence 6357, Ap
41	67.4	97.2	3		Sequence 12381, A
42	67.4	97.2	3		Sequence 6357, Ap
43	67.4	97.2	3		Sequence 12381, A
44	67.4	97.2	3		Sequence 6357, Ap
45	67.4	97.2	3		Sequence 12381, A
46	67.4	97.2	3		Sequence 6357, Ap
47	67.4	97.2	3		Sequence 12381, A
48	67.4	97.2	3		Sequence 6357, Ap
49	67.4	97.2	3		Sequence 12381, A
50	67.4	97.2	3		Sequence 6357, Ap
51	67.4	97.2	3		Sequence 12381, A
52	67.4	97.2	3		Sequence 6357, Ap
53	67.4	97.2	3		Sequence 12381, A
54	67.4	97.2	3		Sequence 6357, Ap
55	67.4	97.2	3		Sequence 12381, A
56	67.4	97.2	3		Sequence 6357, Ap
57	67.4	97.2	3		Sequence 12381, A
58	67.4	97.2	3		Sequence 6357, Ap
59	67.4	97.2	3		Sequence 12381, A
60	67.4	97.2	3		Sequence 6357, Ap
61	67.4	97.2	3		Sequence 12381, A
62	67.4	97.2	3		Sequence 6357, Ap
63	67.4	97.2	3		Sequence 12381, A
64	67.4	97.2	3		Sequence 6357, Ap
65	67.4	97.2	3		Sequence 12381, A
66	67.4	97.2	3		Sequence 6357, Ap
67	67.4	97.2	3		Sequence 12381, A
68	67.4	97.2	3		Sequence 6357, Ap
69	67.4	97.2	3		Sequence 12381, A
70	67.4	97.2	3		Sequence 6357, Ap
71	67.4	97.2	3		Sequence 12381, A
72	67.4	97.2	3		Sequence 6357, Ap
73	67.4	97.2	3		Sequence 12381, A
74	67.4	97.2	3		Sequence 6357, Ap
75	67.4	97.2	3		Sequence 12381, A
76	67.4	97.2	3		Sequence 6357, Ap
77	67.4	97.2	3		Sequence 12381, A
78	67.4	97.2	3		Sequence 6357, Ap
79	67.4	97.2	3		Sequence 12381, A
80	67.4	97.2	3		Sequence 6357, Ap
81	67.4	97.2	3		Sequence 12381, A
82	67.4	97.2	3		Sequence 6357, Ap
83	67.4	97.2	3		Sequence 12381, A
84	67.4	97.2	3		Sequence 6357, Ap
85	67.4	97.2	3		Sequence 12381, A
86	67.4	97.2	3		Sequence 6357, Ap
87	67.4	97.2	3		Sequence 12381, A
88	67.4	97.2	3		Sequence 6357, Ap
89	67.4	97.2	3		Sequence 12381, A
90	67.4	97.2	3		Sequence 6357, Ap
91	67.4	97.2	3		Sequence 12381, A
92	67.4	97.2	3		Sequence 6357, Ap
93	67.4	97.2	3		Sequence 12381, A
94	67.4	97.2	3		Sequence 6357, Ap
95	67.4	97.2	3		Sequence 12381, A
96	67.4	97.2	3		Sequence 6357, Ap
97	67.4	97.2	3		Sequence 12381, A
98	67.4	97.2	3		Sequence 6357, Ap
99	67.4	97.2	3		Sequence 12381, A
100	67.4	97.2	3		Sequence 6357, Ap

28	31	72.1	894	1	US-08-117-362-4	Sequence 4, Appli
29	31	72.1	894	1	US-08-486-922A-4	Sequence 4, Appli
30	31	72.1	894	1	US-08-486-922A-4	Sequence 4, Appli
31	31	72.1	1009	4	US-09-693-146-4	Sequence 4, Appli
32	31	72.1	1432	3	US-08-781-891-71	Sequence 71, Appli
33	31	72.1	1432	4	US-09-618-166-71	Sequence 71, Appli
34	30	69.8	61	4	US-09-134-001C-4574	Sequence 4574, Ap
35	30	69.8	113	4	US-09-489-039A-10318	Sequence 10318, A
36	30	69.8	300	4	US-09-976-594-7	Sequence 7, Appli
37	30	69.8	319	1	US-08-413-118-119	Sequence 119, Appli
38	30	69.8	319	3	US-08-473-446-119	Sequence 119, Appli
39	30	69.8	459	1	US-08-220-151-12	Sequence 12, Appli
40	30	69.8	459	1	US-08-220-151-14	Sequence 14, Appli
41	30	69.8	459	1	US-08-413-118-12	Sequence 12, Appli
42	30	69.8	459	1	US-08-413-118-14	Sequence 14, Appli
43	30	69.8	459	3	US-08-473-446-12	Sequence 12, Appli
44	30	69.8	459	3	US-08-473-446-14	Sequence 14, Appli
45	30	69.8	459	3	US-09-213-053-6	Sequence 6, Appli
46	30	69.8	474	3	US-08-729-416C-1	Sequence 1, Appli
47	30	69.8	474	3	US-08-729-416C-11	Sequence 11, Appli
48	30	69.8	474	3	US-08-807-342B-2	Sequence 2, Appli
49	30	69.8	474	4	US-09-433-353-1	Sequence 1, Appli
50	30	69.8	474	4	US-09-433-353-11	Sequence 11, Appli
51	30	69.8	591	3	US-09-352-159-44	Sequence 44, Appli
52	30	69.8	591	3	US-09-352-159-46	Sequence 46, Appli
53	30	69.8	594	3	US-08-729-416C-7	Sequence 7, Appli
54	30	69.8	594	3	US-08-729-416C-17	Sequence 17, Appli
55	30	69.8	594	4	US-09-433-353-7	Sequence 7, Appli
56	30	69.8	594	4	US-09-433-353-17	Sequence 17, Appli
57	30	69.8	598	3	US-09-352-159-40	Sequence 40, Appli
58	30	69.8	598	3	US-09-352-159-42	Sequence 42, Appli
59	30	69.8	600	3	US-09-352-159-23	Sequence 23, Appli
60	30	69.8	600	3	US-09-352-159-36	Sequence 36, Appli
61	30	69.8	600	3	US-09-352-159-38	Sequence 38, Appli
62	30	69.8	600	3	US-09-352-168-23	Sequence 23, Appli
63	30	69.8	817	4	US-09-489-039A-8184	Sequence 8184, Ap
64	30	69.8	829	3	US-09-352-159-33	Sequence 33, Appli
65	30	69.8	829	3	US-09-352-168-33	Sequence 33, Appli
66	30	69.8	860	1	US-08-117-362-3	Sequence 3, Appli
67	30	69.8	860	1	US-08-486-924-3	Sequence 3, Appli
68	30	69.8	860	4	US-08-486-929A-3	Sequence 3, Appli
69	30	69.8	998	4	US-09-198-452A-841	Sequence 841, App
70	30	69.8	1394	3	US-09-213-053-2	Sequence 2, Appli
71	29	67.4	9	1	US-08-408-604A-155	Sequence 155, App
72	29	67.4	89	4	US-09-530-378A-9	Sequence 9, Appli
73	29	67.4	103	4	US-09-134-001C-3894	Sequence 3894, Ap
74	29	67.4	118	4	US-09-530-378A-4	Sequence 4, Appli
75	29	67.4	130	2	US-08-726-306A-18	Sequence 18, Appli
76	29	67.4	235	4	US-09-489-039A-12012	Sequence 12012, A
77	29	67.4	242	4	US-09-903-814A-20	Sequence 20, Appli
78	29	67.4	246	1	US-08-233-788A-41	Sequence 41, Appli
79	29	67.4	343	4	US-09-655-908-28	Sequence 28, Appli
80	29	67.4	360	4	US-09-134-001C-4278	Sequence 4278, Ap
81	29	67.4	592	4	US-09-643-597-169	Sequence 169, App
82	29	67.4	592	4	US-09-480-884A-169	Sequence 169, App
83	29	67.4	592	4	US-09-542-615A-169	Sequence 169, App
84	29	67.4	592	4	US-09-606-421B-169	Sequence 169, App
85	29	67.4	791	4	US-09-643-597-170	Sequence 170, App
86	29	67.4	791	4	US-09-480-884A-170	Sequence 170, App
87	29	67.4	791	4	US-09-542-615A-170	Sequence 170, App
88	29	67.4	791	4	US-09-606-421B-170	Sequence 170, App
89	29	67.4	920	4	US-09-643-597-357	Sequence 357, App
90	29	67.4	942	4	US-09-919-172-87	Sequence 87, Appli
91	29	67.4	943	4	US-09-193-562D-32	Sequence 32, Appli
92	29	67.4	943	4	US-09-643-597-161	Sequence 161, App
93	29	67.4	943	4	US-09-480-884A-161	Sequence 161, App
94	29	67.4	943	4	US-09-542-615A-161	Sequence 161, App
95	29	67.4	943	4	US-09-606-421B-161	Sequence 161, App
96	29	67.4	943	4	US-09-623-624-4	Sequence 4, Appli
97	29	67.4	943	4	US-09-221-107-161	Sequence 161, App
98	29	67.4	972	3	US-08-335-844A-23	Sequence 23, Appli
99	29	67.4	972	4	US-09-129-366-23	Sequence 23, Appli
100	29	67.4	1171	1	US-07-828-788A-6	Sequence 6, Appli

```

101 29 67.4 1171 1 US-08-278-685-2
102 29 67.4 1171 1 US-08-277-721-2
103 29 67.4 1171 1 US-08-277-721-4
104 29 67.4 1171 1 US-08-602-737-4
105 29 67.4 1171 3 US-09-001-982-4
106 29 67.4 1171 5 PCT-US92-11337-6
107 29 67.4 1186 1 US-08-602-737-6
108 29 67.4 1186 3 US-09-001-982-6
109 29 67.4 1186 3 US-09-171-410-1
110 28 65.1 18 3 US-08-390-353A-8
111 28 65.1 18 3 US-09-101-059-8
112 28 65.1 37 1 US-08-278-091-4
113 28 65.1 37 1 US-08-483-859-4
114 28 65.1 37 2 US-08-472-173-4
115 28 65.1 37 2 US-08-470-419-26
116 28 65.1 37 2 US-08-487-167-4
117 28 65.1 37 2 US-08-761-828-26
118 28 65.1 37 2 US-08-482-816-4
119 28 65.1 37 2 US-08-296-149-4
120 28 65.1 37 2 US-08-390-105-26
121 28 65.1 37 2 US-08-801-499-4
122 28 65.1 37 2 US-08-615-271-4
123 28 65.1 37 3 US-09-074-660-4
124 28 65.1 37 3 US-08-776-949-26
125 28 65.1 37 3 US-09-074-659-4
126 28 65.1 37 3 US-08-482-810-26
127 28 65.1 37 3 US-09-106-468-4
128 28 65.1 37 3 US-09-106-466A-4
129 28 65.1 37 3 US-09-106-467-4
130 28 65.1 37 3 US-09-027-955-26
131 28 65.1 37 4 US-09-636-805-26
132 28 65.1 37 4 US-09-258-128-26
133 28 65.1 37 4 US-09-635-754-26
134 28 65.1 37 4 US-08-680-525-26
135 28 65.1 37 4 US-09-636-223-26
136 28 65.1 220 4 US-09-543-681A-4545
137 28 65.1 221 3 US-08-390-353A-6
138 28 65.1 221 4 US-09-101-059-6
139 28 65.1 228 4 US-09-543-681A-7646
140 28 65.1 269 1 US-07-706-691G-22
141 28 65.1 269 1 US-07-706-691G-26
142 28 65.1 269 1 US-08-254-021-22
143 28 65.1 269 1 US-08-254-021-26
144 28 65.1 269 2 US-08-618-446-22
145 28 65.1 269 2 US-08-618-446-26
146 28 65.1 269 3 US-08-980-135-22
147 28 65.1 269 3 US-08-980-135-26
148 28 65.1 269 4 US-09-585-798-22
149 28 65.1 269 4 US-09-585-798-26
150 28 65.1 287 4 US-09-107-532A-5182

```

ALIGNMENTS

```

RESULT 1
US-09-347-926-18
; Sequence 18, Application US/09347926
; Patent No. 6440386
; GENERAL INFORMATION:
; APPLICANT: LEUNG, SHUI-ON
; TITLE OF INVENTION: STABILIZED RADIOPHOSPHATE-LABELED PROTEINS
; FILE REFERENCE: 018733/0936
; CURRENT APPLICATION NUMBER: US/09/347,926
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 12
; ORGANISM: Homo sapiens
US-09-347-926-18

```

```

Query Match 100.0%; Score 43; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0097;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYNNVQN 8
Db 5 PSYNNVQN 12

```

RESULT 2

```

US-08-353-550-3
; Sequence 3, Application US/08353550
; Patent No. 5744313
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,550
FILING DATE: 09-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-550-3

```

```

Query Match 100.0%; Score 43; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PSYNNVQN 8
Db 306 PSYNNVQN 313

```

RESULT 3

```

US-08-551-687-3
; Sequence 3, Application US/08551687
; Patent No. 5925547
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5925547el Protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000

```

```

; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,687
; FILING DATE: 01-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,550
; FILING DATE: 09-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Matthew B.
; REGISTRATION NUMBER: P39,787
; REFERENCE/DOCKET NUMBER: 2307K-5731
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-687-3
;
; Query Match 100.0%; Score 43; DB 2; Length 423;
; Best Local Similarity 100.0%; Pred. No. 0.52;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 PSYVNVQN 8
Db 306 PSYVNVQN 313
;
; RESULT 4
; US-08-363-215-2
; Sequence 2, Application US/08363215
; Patent No. 5807989
; GENERAL INFORMATION:
; APPLICANT: Margolis, Benjamin L.
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Yajnik, Vijay
; TITLE OF INVENTION: METHODS FOR TREATMENT OR
; TITLE OF INVENTION: DIAGNOSIS OF DISEASES OR
; TITLE OF INVENTION: DISORDERS ASSOCIATED
; TITLE OF INVENTION: WITH AN APB DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,215
; FILING DATE: December 23, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/208
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-363-215-2
;
; Query Match 100.0%; Score 43; DB 1; Length 473;
; Best Local Similarity 100.0%; Pred. No. 0.59;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 PSYVNVQN 8
Db 315 PSYVNVQN 322
;
; RESULT 5
; US-08-807-342B-7
; Sequence 7, Application US/08807342B
; Patent No. 6077686
; GENERAL INFORMATION:
; APPLICANT: Der, Channing
; APPLICANT: O'Bryan, John P.
; APPLICANT: Fawson, Anthony
; TITLE OF INVENTION: No. 6077686e1 SHC Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,342B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,516
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: MTS2USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-807-342B-7

```


Tue May 25 09:20:21 2004

us-09-977-349-5.ra1

```

Query Match      100.0%; Score 43; DB 3; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PSYVNVQN 8
Db      315 PSYVNVQN 322

RESULT 6
US-08-664-962B-4
; Sequence 4, Application US/08664962B
; Patent No. 6218162
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCHANT & GOULD
; STREET: 3100 No. 6218162west Center, 90 South Seventh Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: U.S.A.
; ZIP: 55402-4131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,962B
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Douglas P.
; REGISTRATION NUMBER: 30,300
; REFERENCE/DOCKET NUMBER: MEG 7933.49-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-664-962B-4

Query Match      100.0%; Score 43; DB 3; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PSYVNVQN 8
Db      315 PSYVNVQN 322

RESULT 7
US-09-311-743-4
; Sequence 4, Application US/09311743
; Patent No. 6238903
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/311,743
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7771-32
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-364-7311
; TELEFAX: 416-361-1398
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-311-743-4

Query Match      100.0%; Score 43; DB 3; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PSYVNVQN 8
Db      315 PSYVNVQN 322

RESULT 8
US-08-353-550-8
; Sequence 8, Application US/08353550
; Patent No. 5744313
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,550
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 02307K-057300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-353-550-8

```

Query Match 100.0%; Score 43; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||
Db 316 PSYVNVQN 323

RESULT 9

US-08-551-687-8
; Sequence 8, Application US/08551687
; Patent No. 5925547
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5925547el Protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,687
; FILING DATE: 01-NOV-1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/353,550
; FILING DATE: 09-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Matthew B.
; REGISTRATION NUMBER: P39,787
; REFERENCE/DOCKET NUMBER: 2307K-5731
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-687-8

Query Match 100.0%; Score 43; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||
Db 316 PSYVNVQN 323

RESULT 10

US-08-363-215-1
; Sequence 1, Application US/08363215
; Patent No. 5807989
; GENERAL INFORMATION:
; APPLICANT: Margolis, Benjamin L.
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Yajnik, Vijay
; TITLE OF INVENTION: METHODS FOR TREATMENT OR
; TITLE OF INVENTION: DIAGNOSIS OF DISEASES OR

; TITLE OF INVENTION: DISORDERS ASSOCIATED
; WITH AN APP DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,215
; FILING DATE: December 23, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/208
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-363-215-1

Query Match 97.7%; Score 42; DB 1; Length 469;
Best Local Similarity 97.5%; Pred. No. 0.93;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
|||
Db 311 PSYVNVQN 318

RESULT 11

US-08-807-342B-6
; Sequence 6, Application US/08807342B
; Patent No. 6077686
; GENERAL INFORMATION:
; APPLICANT: Der, Channing
; APPLICANT: O'Bryan, John P.
; APPLICANT: Pawson, Anthony
; TITLE OF INVENTION: No. 6077686el SHC Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

Tue May 25 09:20:21 2004

us-09-977-349-5.ra1

APPLICATION NUMBER: US/08/807,342B
FILING DATE: 28-FEB-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,516
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: MT52USA
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-807-342B-6

Query Match 97.7%; Score 42; DB 3; Length 469;
Best Local Similarity 87.5%; Pred. No. 0.93;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 311 PSYVNVQN 318

RESULT 12
US-08-266-514-27
Sequence 27, Application US/08266514
Patent No. 5594105
GENERAL INFORMATION:
APPLICANT: COMOGLIO, PAOLO
APPLICANT: PONZETTO, CAROLA
TITLE OF INVENTION: PEPTIDE INHIBITORS OF MITOGENESIS AND
TITLE OF INVENTION: MOTOGENESIS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/266,514
FILING DATE: 27-JUN-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9313528
FILING DATE: 30-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9407673.4
FILING DATE: 18-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5594105man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 769-323-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /note= "The Tyr residue at position 5
may be phosphorylated."
US-08-266-514-27

Query Match 86.0%; Score 37; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
Db 3 PSYVNVQ 9

RESULT 13
US-08-654-604-27
Sequence 27, Application US/08654604
Patent No. 5912183
GENERAL INFORMATION:
APPLICANT: COMOGLIO, PAOLO
APPLICANT: PONZETTO, CAROLA
TITLE OF INVENTION: PEPTIDE INHIBITORS OF MITOGENESIS AND
TITLE OF INVENTION: MOTOGENESIS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/654,604
FILING DATE: 29-MAY-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/266,514
FILING DATE: 27-JUN-1994
APPLICATION NUMBER: GB 9313528
FILING DATE: 30-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9407673.4
FILING DATE: 18-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5912183man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 769-323-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5

```
; OTHER INFORMATION: /note= "The Tyr residue at position 5
; OTHER INFORMATION: may be phosphorylated.
US-08-654-604-27
Query Match      86.0%; Score 37; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNVQ 7
DB      3 PSYVNVQ 9

RESULT 14
US-08-408-604A-157
; Sequence 157, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 157:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-408-604A-157
Query Match      81.4%; Score 35; DB 1; Length 8;
Best Local Similarity 87.5%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQ 8
DB      1 PSYVNVQ 8

RESULT 15
US-08-408-604A-156
; Sequence 156, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; Sequence 156, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-408-604A-156
Query Match      81.4%; Score 35; DB 1; Length 9;
Best Local Similarity 87.5%; Pred. No. 3e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQ 8
DB      1 PSYVNVQ 8

RESULT 16
US-08-408-604A-154
; Sequence 154, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 154:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
US-08-408-604A-154

```

```

Query Match      81.4%; Score 35; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 PSYVNQVN 8
Db 6 PSXVNVQN 13

```

```

RESULT 17
US-09-328-352-6758
; Sequence 6758, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6758
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6758

```

```

Query Match      81.4%; Score 35; DB 4; Length 282;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 PSYVNQVN 8
Db 121 PSYFNQVN 128

```

```

RESULT 18
US-09-107-532A-4381
; Sequence 4381, Application US/09107532A
; Patent No. 6563275
; GENERAL INFORMATION:

```

```

;
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4381:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...251
; SEQUENCE DESCRIPTION: SEQ ID NO: 4381:
US-09-107-532A-4381

```

```

Query Match      79.1%; Score 34; DB 4; Length 251;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 PSYVNQV 7
Db 222 PAYVNQV 228

```

```

RESULT 19
US-09-328-352-5843
; Sequence 5843, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5843
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5843

```

Query Match 79.1%; Score 34; DB 4; Length 273;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYVNQV 8
 DB 125 SYXINQ 131

RESULT 20

US-09-621-976-5040
 ; Sequence 5040, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 5040
 ; LENGTH: 121
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-621-976-5040

Query Match 76.7%; Score 33; DB 4; Length 121;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 DB 30 PSYINLQ 36

RESULT 21

US-09-621-976-7453
 ; Sequence 7453, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 7453
 ; LENGTH: 125
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-621-976-7453

Query Match 76.7%; Score 33; DB 4; Length 125;
 Best Local Similarity 71.4%; Pred. No. 15;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
 DB 30 PSYINLQ 36

RESULT 22

US-08-726-306A-37
 ; Sequence 37, Application US/08726306A
 ; Patent No. 5958684
 ; GENERAL INFORMATION:
 ; APPLICANT: van Leeuwen, Frederik Willem

; APPLICANT: Burbach, Johannes Peter Henri
 ; APPLICANT: Grosveld, Franklin G.
 ; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
 ; NUMBER OF SEQUENCES: 189
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner & Witcoff, Ltd.
 ; STREET: 1 Financial Center
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WordPerfect 6.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/726,306A
 ; FILING DATE: 02-Oct-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 95/20080.4
 ; FILING DATE: 02-Oct-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/009,832
 ; FILING DATE: 01-Jan-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Williams, Ph.D., Kathleen M.
 ; REGISTRATION NUMBER: 34,380
 ; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 345-9100
 ; TELEFAX: (617) 345-9111
 ; INFORMATION FOR SEQ ID NO: 37:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 71 amino acids
 ; STRANDEDNESS: single
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 US-08-726-306A-37

Query Match 72.1%; Score 31; DB 2; Length 71;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8
 DB 61 PSYVLLQ 68

RESULT 23

US-09-543-681A-6646
 ; Sequence 6646, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 6646
 ; LENGTH: 274
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 US-09-543-681A-6646

Query Match 72.1%; Score 31; DB 4; Length 274;
 Best Local Similarity 83.3%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 PSYVNV 6
      |||||:
Db      125 PSYVNI 130

RESULT 24
US-09-489-039A-12381
; Sequence 12381, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12381
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12381

Query Match      72.1%; Score 31; DB 4; Length 284;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
      |||||:
Db      135 PSYVNI 140

RESULT 25
US-09-328-352-6357
; Sequence 6357, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6357
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6357

Query Match      72.1%; Score 31; DB 4; Length 285;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
      |||||:
Db      137 PSYVNI 142

RESULT 26
US-09-252-991A-22236
; Sequence 22236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22236
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22236

Query Match      72.1%; Score 31; DB 4; Length 318;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
      |||||:
Db      113 PSYVRVTN 120

RESULT 27
US-09-107-532A-3913
; Sequence 3913, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3913:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...408
; SEQUENCE DESCRIPTION: SEQ ID NO: 3913:
US-09-107-532A-3913

Query Match      72.1%; Score 31; DB 4; Length 408;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
```

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||||:|
Db 98 PSYVTISN 105

RESULT 28

US-08-117-362-4
; Sequence 4, Application US/08117362
; Patent No. 5595872
; GENERAL INFORMATION:
; APPLICANT: Wetterau II, John R.
; APPLICANT: Sharp, Daru Y.
; APPLICANT: Gregg, Richard E.
; TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/117,362
; FILING DATE: 03-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC21b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-117-362-4

Query Match 72.1%; Score 31; DB 1; Length 894;

Best Local Similarity 62.5%; Pred. No. 3.6e+02;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||||:|
Db 552 PSYMDVKN 559

RESULT 29

US-08-486-924-4
; Sequence 4, Application US/08486924
; Patent No. 5783197
; GENERAL INFORMATION:
; APPLICANT: Wetterau II, John R.
; APPLICANT: Sharp, Daru Y.
; APPLICANT: Gregg, Richard E.
; TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey

; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/486,924
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,362
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC21b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-486-924-4

Query Match 72.1%; Score 31; DB 1; Length 894;

Best Local Similarity 62.5%; Pred. No. 3.6e+02;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||||:|
Db 552 PSYMDVKN 559

RESULT 30

US-08-486-929A-4
; Sequence 4, Application US/08486929A
; Patent No. 6492365
; GENERAL INFORMATION:
; APPLICANT: Wetterau II, John R.
; APPLICANT: Sharp, Daru Y.
; APPLICANT: Gregg, Richard E.
; TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,929A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,362
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC21b
; TELECOMMUNICATION INFORMATION:


```

; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-929A-4

Query Match 72.1%; Score 31; DB 4; Length 894;
Best Local Similarity 62.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 552 PSYMDVKN 559

RESULT 31
US-09-693-146-4
; Sequence 4, Application US/09693146
; Patent No. 6413758
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; APPLICANT: Zhu, Zhenyu
; APPLICANT: Xiao, Jian-ping
; TITLE OF INVENTION: Method For Cloning And Expression Of BpmI Restriction
; TITLE OF INVENTION: Endonuclease In E. coli
; FILE REFERENCE: NEB-183
; CURRENT APPLICATION NUMBER: US/09/693,146
; CURRENT FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 1009
; ORGANISM: Bacillus pumilus
; US-09-693-146-4

Query Match 72.1%; Score 31; DB 4; Length 1009;
Best Local Similarity 62.5%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 543 PPYVRQN 550

RESULT 32
US-08-781-891-71
; Sequence 71, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/618,166
; FILING DATE: 17-Jul-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 240052.419C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1432 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 71:

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 609620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-781-891-71

Query Match 72.1%; Score 31; DB 3; Length 1432;
Best Local Similarity 62.5%; Pred. No. 6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
Db 16 PEMNVQN 23

RESULT 33
US-09-618-166-71
; Sequence 71, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/618,166
; FILING DATE: 17-Jul-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 240052.419C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1432 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 71:

```

US-09-618-166-71

Query Match 72.1%; Score 31; DB 4; Length 1432;
Best Local Similarity 62.5%; Pred. No. 6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
| : : : : :
Db 16 PEMVNVQN 23

RESULT 34

US-09-134-001C-4574
; Sequence 4574, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4574
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4574

Query Match 69.8%; Score 30; DB 4; Length 61;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SYVNVQN 8
| : : : : :
Db 53 SYINVSN 59

RESULT 35

US-09-489-039A-10318
; Sequence 10318, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10318
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10318

Query Match 69.8%; Score 30; DB 4; Length 113;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
| : : : : :
Db 102 POLYNISN 109

RESULT 36

US-09-976-594-7

; Sequence 7, Application US/09976594

; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3170236CD1
US-09-976-594-7

Query Match 69.8%; Score 30; DB 4; Length 300;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNV 6
| : : : : :
Db 260 PSYINI 265

RESULT 37

US-08-413-118-119
; Sequence 119, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-413-118-119

Query Match 69.8%; Score 30; DB 1; Length 319;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 235 PGLVNIQN 242

RESULT 38

US-08-473-446-119
; Sequence 119, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-473-446-119

Query Match 69.8%; Score 30; DB 3; Length 319;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 235 PGLVNIQN 242

RESULT 39

US-08-220-151-12
; Sequence 12, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:

; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425086 CURTMS
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-220-151-12

Query Match 69.8%; Score 30; DB 1; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 375 PGLVNIQN 382

RESULT 40

US-08-220-151-14
; Sequence 14, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTWS
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-220-151-14

Query Match 69.8%; Score 30; DB 1; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
DB 375 PGLVNIQN 382

RESULT 41
US-08-413-118-12
; Sequence 12, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-14

Query Match 69.8%; Score 30; DB 1; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
DB 375 PGLVNIQN 382

RESULT 43
US-08-473-446-12
; Sequence 12, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; US-08-413-118-12
```

```
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-473-446-12

Query Match 69.8%; Score 30; DB 3; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 375 PGLVNIQN 382

RESULT 44
US-08-473-446-14
; Sequence 14, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:

; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
; TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200
; CURRENT APPLICATION NUMBER: US/09/213,053
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 9608242
; EARLIER FILING DATE: 1996-06-27
; EARLIER APPLICATION NUMBER: PCT/FR97/01115
; EARLIER FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Canine herpesvirus
; US-09-213-053-6

Query Match 69.8%; Score 30; DB 3; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 375 PGLVNIQN 382

RESULT 46
US-08-729-416C-1
; Sequence 1, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-473-446-14

Query Match 69.8%; Score 30; DB 3; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 375 PGLVNIQN 382

RESULT 45
US-09-213-053-6
; Sequence 6, Application US/09213053
; Patent No. 6159477
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe
; APPLICANT: BAUDU, Philippe
; TITLE OF INVENTION: CANINE HERPESVIRUS BASED RECOMBINANT LIVE VACCINE, IN
; TITLE OF INVENTION: PARTICULAR AGAINST CANINE DISTEMPER, RABIES OR THE
; TITLE OF INVENTION: PARAINFLUENZA 2 VIRUS
; FILE REFERENCE: 454313-2200
; CURRENT APPLICATION NUMBER: US/09/213,053
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 9608242
; EARLIER FILING DATE: 1996-06-27
; EARLIER APPLICATION NUMBER: PCT/FR97/01115
; EARLIER FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Canine herpesvirus
; US-09-213-053-6

Query Match 69.8%; Score 30; DB 3; Length 459;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNVQN 8
Db 375 PGLVNIQN 382
```

ADDRESSER: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1

Query Match 69.8%; Score 30; DB 3; Length 474;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|
DB 302 PTYVNTQ 308

RESULT 47
US-08-729-416C-11
; Sequence 11, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-11

Query Match 69.8%; Score 30; DB 3; Length 474;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|
DB 302 PTYVNTQ 308

RESULT 48
US-08-807-342B-2
; Sequence 2, Application US/08807342B
; Patent No. 6077686
; GENERAL INFORMATION:
; APPLICANT: Der, Channing
; APPLICANT: O'Bryan, John P.
; APPLICANT: Pawson, Anthony
; TITLE OF INVENTION: No. 6077686el SHC Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,342B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,516
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: MTS2USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MOLECULE TYPE: protein
US-08-807-342B-2

Query Match 69.8%; Score 30; DB 3; Length 474;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|
DB 302 PTYVNTQ 308

RESULT 49
US-09-433-353-1
; Sequence 1, Application US/09433353

```
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; TITLE OF INVENTION: AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-1
```

```
Query Match          69.8%; Score 30; DB 4; Length 474;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQ 7
      |:|:|:|
Db      302 PTYVNTQ 308
```

```
RESULT 50
US-09-433-353-11
; Sequence 11, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; TITLE OF INVENTION: AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-433-353-11
```

```
Query Match          69.8%; Score 30; DB 4; Length 474;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQ 7
      |:|:|:|
Db      302 PTYVNTQ 308
```

Search completed: May 24, 2004, 17:33:23
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:31:07 ; Search time 43 Seconds
(without alignments)

51.892 Million cell updates/sec

Title: US-09-977-349-5

Perfect score: 43

Sequence: 1 PSYVNVQN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PTCTUS_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	43	100.0	15	12	US-10-311-510-31	Sequence 31, Appl
2	43	100.0	15	15	US-10-187-967-2	Sequence 2, Appl
3	43	100.0	15	16	US-10-644-055-2	Sequence 2, Appl
4	43	100.0	441	14	US-10-283-881-9	Sequence 9, Appl
5	36	83.7	61	11	US-09-864-408A-5514	Sequence 5514, Ap
6	35	81.4	1131	9	US-09-801-368-72	Sequence 72, Appl
7	35	81.4	1131	15	US-10-369-493-1959	Sequence 1959, Ap
8	34	79.1	471	12	US-10-282-122A-52342	Sequence 52342, A
9	33	76.7	1441	10	US-09-764-891-3838	Sequence 3838, Ap
10	33	76.7	235	12	US-10-425-114-55608	Sequence 55608, A
11	33	76.7	256	15	US-10-104-047-3255	Sequence 3255, Ap
12	33	76.7	267	9	US-09-726-643-63	Sequence 63, Appl
13	33	76.7	267	13	US-10-042-141-63	Sequence 63, Appl
14	33	76.7	401	12	US-10-424-599-156315	Sequence 156315,
15	32	74.4	73	13	US-10-001-835-163	Sequence 163, App

32	74.4	82	12	US-10-001-885-107	Sequence 107, App
32	74.4	82	13	US-10-001-835-218	Sequence 218, App
32	74.4	151	12	US-10-424-599-276996	Sequence 276996, A
32	74.4	415	12	US-10-425-114-62725	Sequence 62725, A
32	74.4	503	12	US-10-282-122A-64241	Sequence 64241, A
32	74.4	578	9	US-09-999-248-2	Sequence 2, Appl
32	74.4	578	12	US-10-418-036-8	Sequence 8, Appl
32	74.4	578	16	US-10-618-839-8	Sequence 8, Appl
32	74.4	622	14	US-10-032-585-7890	Sequence 7890, Ap
32	74.4	678	15	US-10-369-493-12508	Sequence 12508, A
32	74.4	716	15	US-10-369-493-3177	Sequence 3177, Ap
32	74.4	1360	12	US-10-240-315-3	Sequence 3, Appl
32	74.4	1360	16	US-10-433-794-16	Sequence 16, Appl
32	74.4	1360	16	US-10-424-599-261233	Sequence 261233,
31	72.1	39	9	US-09-864-761-44576	Sequence 44576, A
31	72.1	51	9	US-09-864-761-45585	Sequence 45585, A
31	72.1	66	9	US-09-764-847-563	Sequence 563, App
31	72.1	66	14	US-10-092-154-563	Sequence 563, App
31	72.1	273	12	US-10-282-122A-44929	Sequence 44929, A
31	72.1	273	12	US-10-282-122A-50874	Sequence 50874, A
31	72.1	273	12	US-10-282-122A-65381	Sequence 65381, A
31	72.1	273	12	US-10-282-122A-66085	Sequence 66085, A
31	72.1	273	12	US-10-282-122A-73203	Sequence 73203, A
31	72.1	274	12	US-10-282-122A-55812	Sequence 55812, A
31	72.1	274	12	US-10-282-122A-56423	Sequence 56423, A
31	72.1	274	12	US-10-282-122A-67057	Sequence 67057, A
31	72.1	274	12	US-10-282-122A-68564	Sequence 68564, A
31	72.1	274	12	US-10-282-122A-76142	Sequence 76142, A
31	72.1	274	12	US-10-282-122A-77954	Sequence 77954, A
31	72.1	275	12	US-10-282-122A-60013	Sequence 60013, A
31	72.1	276	12	US-10-282-122A-61220	Sequence 61220, A
31	72.1	303	12	US-10-282-122A-58566	Sequence 58566, A
31	72.1	423	12	US-10-425-114-53422	Sequence 53422, A
31	72.1	490	12	US-10-424-599-237538	Sequence 237538,
31	72.1	1432	10	US-09-896-186B-18	Sequence 18, Appl
31	72.1	1432	15	US-10-374-077-71	Sequence 71, Appl
30	69.8	55	12	US-10-424-599-209275	Sequence 209275,
30	69.8	115	12	US-10-424-599-165140	Sequence 165140,
30	69.8	157	12	US-10-424-599-218278	Sequence 218278,
30	69.8	161	12	US-10-424-599-279148	Sequence 279148,
30	69.8	206	12	US-10-424-599-276542	Sequence 276542,
30	69.8	234	12	US-10-424-599-260329	Sequence 260329,
30	69.8	259	12	US-10-424-599-239701	Sequence 239701,
30	69.8	259	12	US-10-424-599-275865	Sequence 275865,
30	69.8	284	10	US-09-769-787-40	Sequence 40, Appl
30	69.8	300	9	US-09-978-295A-464	Sequence 464, App
30	69.8	300	9	US-09-978-697-464	Sequence 464, App
30	69.8	300	9	US-09-978-192A-464	Sequence 464, App
30	69.8	300	9	US-09-999-832A-464	Sequence 464, App
30	69.8	300	10	US-09-978-189-464	Sequence 464, App
30	69.8	300	10	US-09-978-608A-464	Sequence 464, App
30	69.8	300	10	US-09-978-585A-464	Sequence 464, App
30	69.8	300	10	US-09-978-191A-464	Sequence 464, App
30	69.8	300	10	US-09-978-403A-464	Sequence 464, App
30	69.8	300	10	US-09-978-564A-464	Sequence 464, App
30	69.8	300	10	US-09-993-833A-464	Sequence 464, App
30	69.8	300	10	US-09-978-187B-464	Sequence 464, App
30	69.8	300	10	US-09-978-643A-464	Sequence 464, App
30	69.8	300	10	US-09-918-585A-464	Sequence 464, App
30	69.8	300	10	US-09-978-423A-464	Sequence 464, App
30	69.8	300	10	US-09-978-193A-464	Sequence 464, App
30	69.8	300	10	US-09-999-830A-464	Sequence 464, App
30	69.8	300	10	US-09-978-757A-464	Sequence 464, App
30	69.8	300	10	US-09-978-187B-464	Sequence 464, App
30	69.8	300	10	US-09-978-643A-464	Sequence 464, App
30	69.8	300	10	US-09-978-298A-464	Sequence 464, App
30	69.8	300	10	US-09-978-188A-464	Sequence 464, App
30	69.8	300	10	US-09-978-681A-464	Sequence 464, App
30	69.8	300	10	US-09-999-829A-464	Sequence 464, App
30	69.8	300	10	US-09-978-299A-464	Sequence 464, App
30	69.8	300	10	US-09-978-544A-464	Sequence 464, App


```

; GENERAL INFORMATION:
; APPLICANT: BARSANTI, PAUL A
; APPLICANT: BUSSIERE, DIRKSEN
; APPLICANT: HARRISON, STEPHEN D.
; APPLICANT: HEISE, CARLA C.
; APPLICANT: JANSEN, JOHANNA M.
; APPLICANT: JAZAN, ELISA
; APPLICANT: MACHAJEWSKI, TIMOTHY D.
; APPLICANT: MCCBRIDE, CHRISTOPHER
; APPLICANT: MCCREA, WILLIAM R.
; APPLICANT: NG, SIMON
; APPLICANT: NI, ZHI-JTE
; APPLICANT: PECCHI, SABINA
; APPLICANT: PFISTER, KEITH
; APPLICANT: RAMURTHY, SAVITHRI
; APPLICANT: RENHOF, PAUL A.
; APPLICANT: SHAFER, CYNTHIA M.
; APPLICANT: SILVER, JOEL B.
; APPLICANT: WAGMAN, ALLAN
; APPLICANT: WEISMANN, MARION
; TITLE OF INVENTION: BENZIMIDAZOLE QUINOLINONES AND USES THEREOF
; FILE REFERENCE: 072121-0307
; CURRENT APPLICATION NUMBER: US/10/644,055
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 60/484,048
; PRIOR FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 60/478,916
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/460,493
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/460,328
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/460,327
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/428,210
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 60/426,282
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/426,226
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/426,107
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/405,729
; PRIOR FILING DATE: 2002-08-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; OTHER INFORMATION: c-term amidated
US-10-644-055-2

Query Match 100.0%; Score 43; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
DB 7 PSYVNQVN 14

RESULT 4
US-10-283-881-9
; Sequence 9, Application US/10283881
; Publication No. US20030165931A1
; GENERAL INFORMATION:
; APPLICANT: TOCQUE, BRUNO
; APPLICANT: BRACCO, LAURENT

```

```

; APPLICANT: EDON, FLORENCE
; APPLICANT: SCHWEIGHOFFER, FABRIEN
; TITLE OF INVENTION: QUALITATIVE DIFFERENTIAL SCREENING
; FILE REFERENCE: 50146/004003
; CURRENT APPLICATION NUMBER: US/10/283,881
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US 09/623,828
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: PCT/FR99/00547
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/046,920
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 98 02997
; PRIOR FILING DATE: 1998-03-11
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-881-9

Query Match 100.0%; Score 43; DB 14; Length 441;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNQVN 8
DB 315 PSYVNQVN 322

RESULT 5
US-09-864-408A-5514
; Sequence 5514, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5514
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5514

Query Match 83.7%; Score 36; DB 11; Length 61;
Best Local Similarity 85.7%; Pred. No. 7.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNQV 7
DB 54 PSYVNQV 60

RESULT 6
US-09-801-368-72
; Sequence 72, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd

```

```

; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofia
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 1131
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-72

Query Match      81.4%; Score 35; DB 9; Length 1131;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      950 PSYINVLN 957
      |||:|:|
      |||:|:|

RESULT 7
US-10-369-493-1959
; Sequence 1959, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1959
; LENGTH: 1131
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-10-369-493-1959

Query Match      81.4%; Score 35; DB 15; Length 1131;
Best Local Similarity 75.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      950 PSYINVLN 957
      |||:|:|
      |||:|:|

RESULT 8
US-10-282-122A-52342
; Sequence 52342, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

```

```

; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52342
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-10-282-122A-52342

Query Match      79.1%; Score 34; DB 12; Length 471;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 PSYVNVQN 8
Db      114 PSYVNLPN 121
      |||:|:|
      |||:|:|

RESULT 9
US-09-764-891-3838
; Sequence 3838, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3838
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-3838

```

```
Query Match          76.7%; Score 33; DB 10; Length 144;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
   |||:|:|
Db 51 PSYINLQ 57

RESULT 10
US-10-425-114-55608
; Sequence 55608, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55608
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMRONOI.R028D04_FLI.pep
US-10-425-114-55608

Query Match          76.7%; Score 33; DB 12; Length 235;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
   |||||:|
Db 84 PSYVNSRN 91

RESULT 11
US-10-104-047-3255
; Sequence 3255, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: HL-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3255
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3255

Query Match          76.7%; Score 33; DB 15; Length 256;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
   |||||
Db 98 PGYVNVQ 104

RESULT 12
```

```
US-09-726-643-63
; Sequence 63, Application US/09726643
; Patent No. US20020028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: PZ040P1
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-726-643-63

Query Match          76.7%; Score 33; DB 9; Length 267;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
   |||||
Db 69 PGYVNVQ 75

RESULT 13
US-10-042-141-63
; Sequence 63, Application US/10042141
; Publication No. US20020183503A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: PZ040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-141-63

Query Match          76.7%; Score 33; DB 13; Length 267;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
   |||||
Db 69 PGYVNVQ 75

RESULT 14
US-10-424-599-156315
; Sequence 156315, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 156315

; LENGTH: 401

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(401)

; OTHER INFORMATION: unsure at all xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_112173C.1.pap

; US-10-424-599-156315

Query Match 76.7%; Score 33; DB 12; Length 401;

Best Local Similarity 75.0%; Pred. No. 2.1e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8

|||||:

DB 216 PSYVNSKN 223

RESULT 15

US-10-001-835-163

; Sequence 163, Application US/100001835

; Publication No. US20020160387A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Hervé

; APPLICANT: Caifferkey, Robert

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro

; FILE REFERENCE: DEX-0277

; CURRENT APPLICATION NUMBER: US/10/001,835

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/249,997

; PRIOR FILING DATE: 2000-11-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 163

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-001-835-163

Query Match 74.4%; Score 32; DB 13; Length 73;

Best Local Similarity 62.5%; Pred. No. 55;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8

|||||:

DB 13 PSYINKN 20

RESULT 16

US-10-001-885-107

; Sequence 107, Application US/100001885

; Publication No. US20040058319A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Hervé

; APPLICANT: Caifferkey, Robert

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro

; FILE REFERENCE: DEX-0279

; CURRENT APPLICATION NUMBER: US/10/001,885

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/252,061

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: 60/253,257

; PRIOR FILING DATE: 2000-11-27

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 107

; LENGTH: 82

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-001-885-107

Query Match 74.4%; Score 32; DB 12; Length 82;

Best Local Similarity 62.5%; Pred. No. 63;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8

|||||:

DB 13 PSYINKN 20

RESULT 17

US-10-001-835-218

; Sequence 218, Application US/10001835

; Publication No. US20020160387A1

; GENERAL INFORMATION:

; APPLICANT: Salceda, Susana

; APPLICANT: Macina, Roberto

; APPLICANT: Recipon, Hervé

; APPLICANT: Caifferkey, Robert

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro

; FILE REFERENCE: DEX-0277

; CURRENT APPLICATION NUMBER: US/10/001,835

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: 60/249,997

; PRIOR FILING DATE: 2000-11-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 218

; LENGTH: 82

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-001-835-218

Query Match 74.4%; Score 32; DB 13; Length 82;

Best Local Similarity 62.5%; Pred. No. 63;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8

|||||:

DB 13 PSYINKN 20

RESULT 18

US-10-424-599-276996

; Sequence 276996, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 276996

LENGTH: 151
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_92146C.1.pep
US-10-424-599-276996

Query Match 74.4%; Score 32; DB 12; Length 151;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|||||
DB 32 PSYVNVQ 38

RESULT 19
US-10-425-114-62725
Sequence 62725, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 62725
LENGTH: 415
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3592-013-H7_FLI.pep
US-10-425-114-62725

Query Match 74.4%; Score 32; DB 12; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
|||||
DB 387 PSYVNV 392

RESULT 20
US-10-282-122A-64241
Sequence 64241, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Lianguo
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 64241
LENGTH: 503
TYPE: PRT
ORGANISM: Mycoplasma pneumoniae
US-10-282-122A-64241

Query Match 74.4%; Score 32; DB 12; Length 503;
Best Local Similarity 62.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
||:|
DB 343 PSFINEQN 350

RESULT 21
US-09-999-248-2
Sequence 2, Application US/09999248
Patent No. US20020176852A1
GENERAL INFORMATION:
APPLICANT: Lambeth, J. David
APPLICANT: Cheng, Guangjie
TITLE OF INVENTION: Mitogenic Oxygenase Regulators
FILE REFERENCE: 05501-0180 43150-266489
CURRENT APPLICATION NUMBER: US/09/999,248
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/249,305
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/251,364
PRIOR FILING DATE: 2000-12-05
PRIOR APPLICATION NUMBER: US 60/289,172
PRIOR FILING DATE: 2001-05-07
PRIOR APPLICATION NUMBER: US 60/289,537
PRIOR FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 578
TYPE: PRT
ORGANISM: Homo sapiens
US-09-999-248-2

Query Match 74.4%; Score 32; DB 9; Length 578;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|||:
DB 489 PDYVNIQ 495

RESULT 22

```
US-10-418-036-8
; Sequence 8, Application US/10418036
; Publication No. US20030225117A1
; GENERAL INFORMATION:
; APPLICANT: Gronberg, Alvar
; APPLICANT: Wikstrom, Per
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: 13425-110001
; CURRENT APPLICATION NUMBER: US/10/418,036
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: SE 0201152-6
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/410,626
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-036-8

Query Match 74.4%; Score 32; DB 12; Length 578;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|
Db 489 PDYVNIQ 495

RESULT 23
US-10-618-839-8
; Sequence 8, Application US/10618839
; Publication No. US20040093628A1
; GENERAL INFORMATION:
; APPLICANT: Lambeth, J. David
; APPLICANT: Cheng, Guangjie
; APPLICANT: McCoy, James
; TITLE OF INVENTION: Methods and Transgenic Mouse Model for Identifying and Modulating
; TITLE OF INVENTION: Factors Involved in the Production of Reactive Oxygen
; TITLE OF INVENTION: Intermediates
; FILE REFERENCE: 05501-0211 (43150-286808)
; CURRENT APPLICATION NUMBER: US/10/618,839
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US 60/395,498
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-839-8

Query Match 74.4%; Score 32; DB 16; Length 578;
Best Local Similarity 71.4%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 7
|:|:|:|
Db 489 PDYVNIQ 495

RESULT 24
US-10-032-585-7890
; Sequence 7890, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28

US-10-032-585-7890
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7890
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7890

Query Match 74.4%; Score 32; DB 14; Length 622;
Best Local Similarity 62.5%; Pred. No. 5.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8
|:|:|:|
Db 600 PNYLVNN 607

RESULT 25
US-10-369-493-12508
; Sequence 12508, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12508
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ..(678)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12508

Query Match 74.4%; Score 32; DB 15; Length 678;
Best Local Similarity 62.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQ 8
|:|:|:|
Db 531 PTYINVLN 538

RESULT 26
US-10-369-493-3177
; Sequence 3177, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
```

; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 3177
 ; LENGTH: 716
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(716)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-3177

Query Match 74.4%; Score 32; DB 15; Length 716;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYNNV 6
 DB 368 PSYNNV 373

RESULT 27
 US-10-240-315-3
 ; Sequence 3, Application US/10240315
 ; Publication No. US20030224378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FLOWMAN, GREGORY D.
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: MANNING, GERARD
 ; APPLICANT: SUDARSANAM, SUCHA
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: CAENEPEEL, SEAN R.
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
 ; FILE REFERENCE: 038602-1456
 ; CURRENT APPLICATION NUMBER: US/10/240,315
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: PCT/US01/11675
 ; PRIOR FILING DATE: 2001-04-10
 ; PRIOR APPLICATION NUMBER: 60/195,953
 ; PRIOR FILING DATE: 2000-04-10
 ; PRIOR APPLICATION NUMBER: 60/201,015
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: 60/213,805
 ; PRIOR FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 1360
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-240-315-3

Query Match 74.4%; Score 32; DB 12; Length 1360;
 Best Local Similarity 62.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYNNV 8
 DB 589 PSYNNV 596

RESULT 28
 US-10-433-794-16
 ; Sequence 16, Application US/10433794
 ; Publication No. US2004007704A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YUE, Henry; DING, Li;
 ; APPLICANT: LAL, Preeti G.; GRIFFIN, Jennifer A.;
 ; APPLICANT: GURURAJAN, Rajagopal; BAUGHN, Mariah R.;
 ; APPLICANT: ISON, Craig H.; RAMKUNAR, Jayalaxmi;
 ; APPLICANT: TRIBOULEY, Catherine M.; SWARNAKAR, Anita;

; APPLICANT: BURFORD, Neil; BANDMAN, Olga;
 ; APPLICANT: THORNTON, Michael; KHAN, Farrah A.;
 ; APPLICANT: WALIA, Navinder K.; NGUYEN, Daniel B.;
 ; APPLICANT: ELLIOTT, Vicki S.; XU, Yuming;
 ; APPLICANT: LU, Yan; HAFALIA, April J.A.;
 ; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
 ; APPLICANT: ARVIZU, Chandra S.; FORSYTHE, Ian J.
 ; TITLE OF INVENTION: KINASES AND PHOSPHATASES
 ; FILE REFERENCE: PI-0311 USN
 ; CURRENT APPLICATION NUMBER: US/10/433,794
 ; CURRENT FILING DATE: 2003-06-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/47431
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/254,034
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: US 60/251,814
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/255,756
 ; PRIOR FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: US 60/256,172
 ; PRIOR FILING DATE: 2000-12-15
 ; PRIOR APPLICATION NUMBER: US 60/257,416
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 60/260,912
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: US 60/264,644
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: US 60/266,017
 ; PRIOR FILING DATE: 2001-02-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 16
 ; LENGTH: 1360
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 55064363CD1
 US-10-433-794-16

Query Match 74.4%; Score 32; DB 16; Length 1360;
 Best Local Similarity 62.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYNNV 8
 DB 589 PSYNNV 596

RESULT 29
 US-10-424-599-261233
 ; Sequence 261233, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 261233
 ; LENGTH: 35
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(35)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_77916C.1.1.pap
US-10-424-599-261233

Query Match 72.1%; Score 31; DB 12; Length 35;
Best Local Similarity 62.5%; Pred. No. 41;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
DB 7 POYVNLQS 14

RESULT 30
US-09-864-761-44576
; Sequence 44576, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44576
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006385.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.76
; OTHER INFORMATION: EST_HUMAN HIT: AA760692.1, EVALUE 5.00e-17
US-09-864-761-44576

Query Match 72.1%; Score 31; DB 9; Length 39;
Best Local Similarity 75.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
DB 15 PSYVLLQN 22

RESULT 31
US-09-864-761-45585
; Sequence 45585, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45585
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008958.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.48
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
; OTHER INFORMATION: SWISSPROT HIT: Q57984, EVALUE 1.60e+00

; OTHER INFORMATION: EST_HUMAN HIT: AW207853.1, EVALUE 3.80e+00
US-09-864-761-45585

Query Match 72.1%; Score 31; DB 9; Length 51;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PSYVNVQN 8
| | | | |
DB 3 PSYNEQN 10

RESULT 32

US-09-764-847-563
; Sequence 563, Application US/05764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 563

; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-563

Query Match 72.1%; Score 31; DB 9; Length 66;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YVNVQN 8
| | | | |
DB 22 YVNVQN 27

RESULT 33

US-10-092-154-563
; Sequence 563, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 563

; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (66)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-563

Query Match 72.1%; Score 31; DB 14; Length 66;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YVNVQN 8
| | | | |
DB 22 YVNVQN 27

RESULT 34

US-10-282-122A-44929
; Sequence 44929, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44929
; LENGTH: 273

```

; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44929

Query Match          72.1%; Score 31; DB 12; Length 273;
Best Local Similarity 83.3%; Pred.No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
Db      125 PSYVNI 130
      |||||:

RESULT 35
US-10-282-122A-50874
; Sequence 50874, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50874
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50874

Query Match          72.1%; Score 31; DB 12; Length 273;
Best Local Similarity 83.3%; Pred.No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
Db      126 PSYVNI 131
      |||||:

RESULT 36
US-10-282-122A-65381
; Sequence 65381, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 65381
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-65381

Query Match          72.1%; Score 31; DB 12; Length 273;
Best Local Similarity 83.3%; Pred.No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYVNV 6
Db      125 PSYVNI 130
      |||||:

RESULT 37
US-10-282-122A-66085
; Sequence 66085, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

```

```
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60085
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-10-282-122A-66085
```

```
Query Match 72.1%; Score 31; DB 12; Length 273;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 PSYNNV 6
Db 125 PSYNNI 130
```

```
RESULT 38
US-10-282-122A-73203
; Sequence 73203, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73203
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-73203
```

```
Query Match 72.1%; Score 31; DB 12; Length 273;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 PSYNNV 6
Db 125 PSYNNI 130
```

```
RESULT 39
US-10-282-122A-55812
; Sequence 55812, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55812
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55812

Query Match 72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 125 PSYVNI 130

RESULT 40

US-10-282-122A-56423
; Sequence 56423, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56423
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Escherichia coli

US-10-282-122A-56423

Query Match 72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 125 PSYVNI 130

RESULT 41

US-10-282-122A-67057
; Sequence 67057, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67057
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Pasteurella multocida

US-10-282-122A-67057

Query Match 72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 124 PSYVNI 129

RESULT 42

US-10-282-122A-68564
; Sequence 68564, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel

Query Match 72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6
Db 125 PSYVNI 130

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68564
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68564

Query Match      72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYNNV 6
DB      125 PSYNNI 130

RESULT 43
US-10-282-122A-76142
; Sequence 76142, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

```

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76142
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-76142

Query Match      72.1%; Score 31; DB 12; Length 274;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PSYNNV 6
DB      125 PSYNNI 130

RESULT 44
US-10-282-122A-77954
; Sequence 77954, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 77954
; LENGTH: 274

; TYPE: PRT

; ORGANISM: Yersinia pestis

US-10-282-122A-77954

Query Match 72.1%; Score 31; DB 12; Length 274;

Best Local Similarity 83.3%; Pred. No. 3.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6

Db 125 PSYVNI 130

RESULT 45

US-10-282-122A-60013

; Sequence 60013, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60013

; LENGTH: 275

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-10-282-122A-60013

Query Match 72.1%; Score 31; DB 12; Length 275;

Best Local Similarity 83.3%; Pred. No. 3.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6

Db 125 PSYVNI 130

RESULT 46

US-10-282-122A-61220

; Sequence 61220, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 61220

; LENGTH: 276

; TYPE: PRT

; ORGANISM: Legionella pneumophila

US-10-282-122A-61220

Query Match 72.1%; Score 31; DB 12; Length 276;

Best Local Similarity 83.3%; Pred. No. 3.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSYVNV 6

Db 125 PSYVNI 130

RESULT 47

US-10-282-122A-58566

; Sequence 58566, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

```

; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58566
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-282-122A-58566

Query Match          72.1%; Score 31; DB 12; Length 303;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSYVNV 6
Db 153 PSYVNI 158

RESULT 48
US-10-425-114-53422
; Sequence 53422, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53422
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700891859_FLI.pcp

```

```

US-10-425-114-53422

Query Match          72.1%; Score 31; DB 12; Length 423;
Best Local Similarity 62.5%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNV 8
Db 264 PKYVNI 271

RESULT 49
US-10-424-599-237538
; Sequence 237538, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 237538
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_56521C.1.pcp
US-10-424-599-237538

Query Match          72.1%; Score 31; DB 12; Length 490;
Best Local Similarity 62.5%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSYVNV 8
Db 331 PKYVNI 338

RESULT 50
US-09-896-186B-18
; Sequence 18, Application US/09896186B
; Publication No. US20030166227A1
; GENERAL INFORMATION:
; APPLICANT: Joshua Z. Levin
; APPLICANT: Ken Phillips
; APPLICANT: Greg Budziszewski
; APPLICANT: Fred Meins
; APPLICANT: Zhenya Glazov
; TITLE OF INVENTION: Methods of Controlling Gene Expression
; FILE REFERENCE: PB/5-31481A
; CURRENT APPLICATION NUMBER: US/09/896,186B
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-186B-18

Query Match          72.1%; Score 31; DB 10; Length 1432;
Best Local Similarity 62.5%; Pred. No. 2e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSYVNV 8
Db 16 PEMVNV 23

```


Search completed: May 24, 2004, 17:34:20
Job time : 46 secs

```

QY      2 RQPKIWFNNRRKPKWK 17
      ||||| |||||
Db      48 RQVKIWFQNNRRKWK 63

RESULT 16
Q63255 PRELIMINARY; PRT; 85 AA.
AC Q63255;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hox1.4 protein (Fragment).
GN HOX1.4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RX MEDLINE=94347374; PubMed=7915120;
RA Gorski D.H., LePage D.F., Walsh K.;
RT "Cloning and sequence analysis of homeobox transcription factor cDNAs
RT with an inosine-containing probe.";
RL Biotechniques 16:856-858(1994).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; L03557; AAA67845.1; -.
DR PIR; I65198; I65198.
DR HSP; FI4653; I672.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
RP SEQUENCE 85 AA; 9957 MW; 18F8AF9BDC69A09 CRC64;

Query Match 66.0%; Score 68; DB 11; Length 85;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17
      ||||| |||||
Db      22 RQVKIWFQNNRRKWK 37

RESULT 17
Q61681 PRELIMINARY; PRT; 85 AA.
AC Q61681;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Homeobox protein (Hox1-3) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC MEDLINE=87004567; PubMed=3019676;
RA Duboule D., Baron A.J., Maehl P., Galliot B.;
RT "A new homeo-box is present in overlapping cosmid clones which define
RT the mouse HOX-1 locus.";
RL EMBO J. 5:1973-1980(1986).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

```

```

DR EMBL; M26802; AAA37841.1; -.
DR PIR; A25472; A25472.
DR HSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
RP SEQUENCE 85 AA; 10431 MW; 27E7C7001242AC78 CRC64;

Query Match 66.0%; Score 68; DB 11; Length 85;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17
      ||||| |||||
Db      52 RQVKIWFQNNRRKWK 67

RESULT 18
OS7360 PRELIMINARY; PRT; 86 AA.
ID OS7360
AC OS7360;
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hoxa8 protein (Fragment).
GN Hoxa8 OR HOXA8.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Prince V.E., Joly L., Ekker M., Ho R.K.;
RT "Zebrafish hox genes: genomic organization and modified colinear
RT expression patterns in the trunk.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; Y14534; CAA74869.1; -.
DR HSP; P02833; 1SAN.
DR ZFIN; ZDB-GENE-980526-291; hoxb8b.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
RP SEQUENCE 86 AA; 10335 MW; 1C9226EE0B5DBE2B CRC64;

Query Match 66.0%; Score 68; DB 13; Length 86;
Best Local Similarity 75.0%; Pred. No. 0.006;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17

```



```

DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nucleur protein.
FT NON_TER 1 1
FT NON_TER 103 103
SQ SEQUENCE 103 AA; 12388 MW; A3B8A296AC02ACD4 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 103;
Best Local Similarity 75.0%; Pred. No. 0.0071;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
   |||||
Db 80 RQVKIWFQNRKMKWK 95

RESULT 22
Q9PMM1 PRELIMINARY; PRT; 104 AA.
AC Q9PMM1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox protein (Fragment).
GN HOXB8B.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99051425; PubMed=9831563;
RA Amores A., Force A., Yan Y.-L., Wang Y.-L., Fritz A., Prince V.,
RA Ho R., Amemiya C., Langeland J., Westerfield M., Ekker M.,
RA Postlethwait J.;
RT "Zebrafish box clusters and vertebrate genome evolution.";
RL Science 282:1711-1714(1998).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF071255; AAD15948.1; -.
DR HSP; P02833; 9ANT.
DR TRANSFAC; T03636; -.
DR ZFIN; ZDB-GENE-980526-291; hoxb8b.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000047; HTH_lambrepresr.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nucleur protein.
FT NON_TER 1 1
SQ SEQUENCE 104 AA; 12587 MW; E42C7B28D0E13AB6 CRC64;

Query Match 66.0%; Score 68; DB 13; Length 104;
Best Local Similarity 75.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
   |||||
Db 44 RQVKIWFQNRKMKWK 59

RESULT 23
Q801D2 PRELIMINARY; PRT; 106 AA.
AC Q801D2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HoxA4 (Fragment).
OS Latimeria menadoensis (Indonesian coelacanth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OX NCBI_TaxID=106881;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22457206; PubMed=12547909;
RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RT "Hox gene clusters in the Indonesian coelacanth, Latimeria
RT menadoensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1084-1088(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Koh E.G.L., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY183725; AAC43018.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
FT NON_TER 1 1
SQ SEQUENCE 106 AA; 12712 MW; 9F38476519F956CD CRC64;

Query Match 66.0%; Score 68; DB 13; Length 106;
Best Local Similarity 75.0%; Pred. No. 0.0073;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFNRRKPKWK 17
   |||||
Db 51 RQVKIWFQNRKMKWK 66

RESULT 24
Q96896 PRELIMINARY; PRT; 125 AA.
AC Q96896;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sex comb reduced (Fragment).
GN SCR.
OS Sacculina carcini.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;
OC Rhizocephala; Kentronogida; Sacculinidae; Sacculina.
OX NCBI_TaxID=51650;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98334561; PubMed=9667986;
RA Mouchel-Vielh E., Rigolot C., Gilbert J.M., Deutsch J.S.;
RT "Molecules and the body plan: the Hox genes of Cirripedes
RT (Crustacea).";
RL Mol. Phylogenet. Evol. 9:382-389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2199721; PubMed=12004965;
RA Mouchel-Vielh E., Blin M., Rigolot C., Deutsch J.S.;
RT "Expression of a homologue of the fushi tarazu (ftz) gene in a
RT cirripede crustacean.";
RL Evol. Dev. 4:76-85(2002).
DR EMBL; AF393441; AAM50457.1; -.
DR HSP; P02833; 1SAN.
DR GO; GO:0005634; C:nucleus; IEA.

```

Job time : 31.4054 secs

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
FT NON TER 1
SQ SEQUENCE 125 AA; 14906 MW; 49711D87C57AC10C CRC64;

Query Match 66.0%; Score 68; DB 5; Length 125;
Best Local Similarity 75.0%; Pred. No. 0.0086;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWPENRRKPWK 17
||| ||| ||| |||
DB 50 RQVKIWFQNRMRKWK 65

RESULT 25

Q8BQA3 PRELIMINARY; PRT; 141 AA.
AC Q8BQA3
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Homeo box C8 (Fragment).
GN HOXC8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK051173; BAC34544.1; -.
DR MGD; MGI-96198; Hoxc8.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH lambdarepressor.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
FT NON TER 1
SQ SEQUENCE 141 AA; 16700 MW; 7387B143F4514FAC CRC64;

Query Match 66.0%; Score 68; DB 11; Length 141;
Best Local Similarity 75.0%; Pred. No. 0.0097;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWPENRRKPWK 17
||| ||| ||| |||
DB 90 RQVKIWFQNRMRKWK 105

Search completed: May 24, 2004, 17:24:40

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:22:27 ; Search time 12.1757 Seconds
(without alignments)
72.082 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IROPKIWFPNRRKPKWK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	96.1	16	3	US-08-849-486-6
2	75	72.8	16	3	US-08-849-486-5
3	70	68.0	22	4	US-09-155-165-11
4	68	66.0	61	3	US-08-751-344B-7
5	67	65.0	16	2	US-08-928-958-7
6	67	65.0	16	2	US-08-810-540-3
7	67	65.0	16	2	US-08-810-540-6
8	67	65.0	16	2	US-09-072-429-7
9	67	65.0	16	3	US-08-964-302A-6
10	67	65.0	16	3	US-09-116-294-4
11	67	65.0	16	3	US-08-964-614A-4
12	67	65.0	16	3	US-08-849-486-1
13	67	65.0	16	3	US-08-849-486-4
14	67	65.0	16	3	US-09-208-966-54
15	67	65.0	16	3	US-09-308-935-8
16	67	65.0	16	3	US-09-443-416A-6
17	67	65.0	16	4	US-09-296-089-33
18	67	65.0	16	4	US-09-419-826-35
19	67	65.0	16	4	US-09-302-305C-10
20	67	65.0	16	4	US-09-346-847-1
21	67	65.0	16	4	US-09-346-847-25
22	67	65.0	16	4	US-09-057-363C-47
23	67	65.0	16	4	US-09-043-560B-3
24	67	65.0	16	4	US-09-648-400A-29
25	67	65.0	16	4	US-09-227-652B-4
26	67	65.0	16	4	US-09-780-070-38
27	67	65.0	16	4	US-08-610-220B-9

28	67	65.0	16	4	US-09-775-052A-54	Sequence 54, Appl
29	67	65.0	16	4	US-09-155-165-22	Sequence 22, Appl
30	67	65.0	16	4	US-09-792-480-29	Sequence 29, Appl
31	67	65.0	16	4	US-09-551-976-33	Sequence 33, Appl
32	67	65.0	16	4	US-09-265-107-47	Sequence 47, Appl
33	67	65.0	17	4	US-09-346-847-17	Sequence 17, Appl
34	67	65.0	17	4	US-09-346-847-20	Sequence 20, Appl
35	67	65.0	17	4	US-09-346-847-22	Sequence 22, Appl
36	67	65.0	17	4	US-09-346-847-27	Sequence 27, Appl
37	67	65.0	17	4	US-09-648-400A-30	Sequence 30, Appl
38	67	65.0	18	3	US-08-838-545-20	Sequence 20, Appl
39	67	65.0	18	3	US-09-349-532-20	Sequence 20, Appl
40	67	65.0	19	4	US-09-346-847-23	Sequence 23, Appl
41	67	65.0	19	4	US-09-658-517C-7	Sequence 7, Appl
42	67	65.0	20	4	US-09-346-847-16	Sequence 16, Appl
43	67	65.0	20	4	US-09-346-847-18	Sequence 18, Appl
44	67	65.0	20	4	US-09-346-847-30	Sequence 30, Appl
45	67	65.0	20	4	US-09-658-517C-8	Sequence 8, Appl
46	67	65.0	21	4	US-08-610-220B-11	Sequence 11, Appl
47	67	65.0	22	4	US-09-346-847-28	Sequence 28, Appl
48	67	65.0	22	4	US-09-057-363C-50	Sequence 50, Appl
49	67	65.0	22	4	US-08-610-220B-10	Sequence 10, Appl
50	67	65.0	22	4	US-09-155-165-5	Sequence 5, Appl
51	67	65.0	22	4	US-09-265-107-50	Sequence 50, Appl
52	67	65.0	24	4	US-09-419-826-34	Sequence 34, Appl
53	67	65.0	24	4	US-09-428-082B-332	Sequence 332, App
54	67	65.0	27	3	US-09-051-934-51	Sequence 51, Appl
55	67	65.0	27	3	US-09-051-934-52	Sequence 52, Appl
56	67	65.0	27	4	US-09-040-725A-2	Sequence 2, Appl
57	67	65.0	34	4	US-09-347-504-79	Sequence 79, Appl
58	67	65.0	34	4	US-10-161-499-79	Sequence 79, Appl
59	67	65.0	36	4	US-09-428-082B-331	Sequence 331, App
60	67	65.0	61	2	US-08-202-044-3	Sequence 3, Appl
61	67	65.0	61	3	US-08-751-344B-3	Sequence 3, Appl
62	67	65.0	61	3	US-08-751-344B-6	Sequence 6, Appl
63	67	65.0	61	3	US-08-751-344B-9	Sequence 9, Appl
64	66	64.1	16	4	US-09-466-772-7	Sequence 7, Appl
65	66	64.1	20	4	US-09-466-772-3	Sequence 3, Appl
66	66	64.1	21	4	US-09-466-772-1	Sequence 1, Appl
67	66	64.1	22	4	US-09-466-772-2	Sequence 2, Appl
68	66	64.1	23	4	US-09-466-772-4	Sequence 4, Appl
69	64	62.1	16	3	US-08-849-486-7	Sequence 7, Appl
70	64	62.1	302	2	US-08-203-532F-4	Sequence 4, Appl
71	64	62.1	302	3	US-08-950-860-16	Sequence 16, Appl
72	64	62.1	302	3	US-09-078-465-4	Sequence 4, Appl
73	64	62.1	302	5	PCT-US95-01882A-4	Sequence 4, Appl
74	64	62.1	303	2	US-08-203-532F-2	Sequence 2, Appl
75	64	62.1	303	3	US-09-078-465-2	Sequence 2, Appl
76	64	62.1	303	5	PCT-US95-01882A-2	Sequence 2, Appl
77	62	60.2	15	2	US-08-810-540-4	Sequence 4, Appl
78	62	60.2	16	4	US-09-792-480-30	Sequence 30, Appl
79	62	60.2	42	3	US-08-751-344B-4	Sequence 4, Appl
80	62	60.2	283	1	US-08-583-672-2	Sequence 2, Appl
81	62	60.2	283	2	US-08-202-044-2	Sequence 2, Appl
82	62	60.2	283	3	US-08-751-344B-2	Sequence 2, Appl
83	62	60.2	284	2	US-08-320-148B-2	Sequence 2, Appl
84	62	60.2	284	3	US-08-589-028-6	Sequence 6, Appl
85	62	60.2	284	3	US-08-784-582-6	Sequence 6, Appl
86	62	60.2	284	3	US-08-785-271-6	Sequence 2, Appl
87	62	60.2	284	3	US-09-031-898-2	Sequence 2, Appl
88	60	58.3	16	3	US-08-849-486-8	Sequence 8, Appl
89	60	58.3	16	3	US-09-208-966-1	Sequence 1, Appl
90	60	58.3	16	4	US-09-775-052A-1	Sequence 1, Appl
91	60	58.3	300	4	US-09-162-524-1	Sequence 1, Appl
92	60	58.3	349	4	US-09-162-524-3	Sequence 3, Appl
93	59	57.3	367	4	US-09-009-816-2	Sequence 2, Appl
94	58	56.3	16	4	US-09-296-089-34	Sequence 34, Appl
95	58	56.3	16	4	US-09-057-363C-48	Sequence 48, Appl
96	58	56.3	16	4	US-09-551-976-34	Sequence 34, Appl
97	58	56.3	16	4	US-09-265-107-48	Sequence 48, Appl
98	58	56.3	22	4	US-09-265-107-72	Sequence 72, Appl
99	57	55.3	16	4	US-09-402-929-7	Sequence 7, Appl
100	56	54.4	15	4	US-10-083-889-16	Sequence 16, Appl

ALIGNMENTS

```

RESULT 1
US-08-849-486-6
; Sequence 6, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-849-486-6

Query Match          96.1%; Score 99; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17
        |||
Db       1 RQPKIWFNNRRKPKWK 16

RESULT 2
US-08-849-486-5
; Sequence 5, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-849-486-5

Query Match          96.1%; Score 99; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17
        |||
Db       1 RQPKIWFNNRRKPKWK 16

```

```

Query Match          72.8%; Score 75; DB 3; Length 16;
Best Local Similarity 81.2%; Pred. No. 4e-05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 RQPKIWFNNRRKPKWK 17
        |||
Db       1 RQPKIWFNNRRKPKWK 16

RESULT 3
US-09-155-165-11
; Sequence 11, Application US/09155165
; Patent No. 6660830
; GENERAL INFORMATION:
; APPLICANT: Radulescu, Razvan T
; TITLE OF INVENTION: PEPTIDES WITH ANTIPROLIFERATIVE PROPERTIES
; FILE REFERENCE: 201196/20
; CURRENT APPLICATION NUMBER: US/09/155,165
; CURRENT FILING DATE: 1999-06-07
; PRIOR APPLICATION NUMBER: 09/155,165
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: PCT/DE97/00643
; PRIOR FILING DATE: 1997-03-26
; PRIOR APPLICATION NUMBER: DE 196 11 939.1
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: DE 196 53 445.3
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 22
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide
; NAME/KEY: UNSURE
; LOCATION: (1)..(22)
; OTHER INFORMATION: Where all amino acids may be in L or D
; OTHER INFORMATION: configuration
; OTHER INFORMATION:
US-09-155-165-11

Query Match          68.0%; Score 70; DB 4; Length 22;
Best Local Similarity 70.6%; Pred. No. 0.00031;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 IRQPKIWFNNRRKPKWK 17
        :|||
Db       6 VRQIKIWFNNRRKPKWK 22

RESULT 4
US-08-751-344B-7
; Sequence 7, Application US/08751344B
; Patent No. 6210960
; GENERAL INFORMATION:
; APPLICANT: Habener M.D., Joel P.
; APPLICANT: Miller Ph.D., Christopher P.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1

```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,344B
; FILING DATE: 19-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/202,044
; FILING DATE: 23-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-751-344B-7

Query Match 66.0%; Score 68; DB 3; Length 61;
Best Local Similarity 75.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RPKIWFNNRRKPKWK 17
Db 44 RQVKIWFQNNRRMKWK 59

RESULT 5
US-08-928-958-7
; Sequence 7, Application US/08928958
; Patent No. 5877282
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CLEVELAND, JEFFREY S.
; APPLICANT: BLAKE, JAMES
; APPLICANT: HAFAR, OMAR K.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,958
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026978
; FILING DATE: 20-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0019
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-928-958-7

Query Match 65.0%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RPKIWFNNRRKPKWK 17
Db 1 RQIKIWFQNNRRMKWK 16

RESULT 6
US-08-810-540-3
; Sequence 3, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
; APPLICANT: Troy, Carol M.
; APPLICANT: Shelanski, Michael L.
; TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,540
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Reg., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0526
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-810-540-3

Query Match 65.0%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RPKIWFNNRRKPKWK 17
Db 1 RQIKIWFQNNRRMKWK 16

RESULT 7
US-08-810-540-6
; Sequence 6, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
```


APPLICANT: Troy, Carol M.
APPLICANT: Shelanski, Michael L.
TITLE OF INVENTION: ANTISENSE COMPOUNDS WHICH PREVENT CELL
DEATH AND USES THEREOF
TITLE OF INVENTION: 7
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,540
FILING DATE: 03-MAR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/51247
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-810-540-6

Query Match 65.0%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RPKIWFENRRKPKWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 8
US-09-072-429-7
Sequence 7, Application US/09072429
Patent No. 5962415
GENERAL INFORMATION:
APPLICANT: Nadler, Steven G.
TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,429
FILING DATE: 04-MAY-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Klein, Christopher A.
REGISTRATION NUMBER: 34,363
REFERENCE/DOCKET NUMBER: ON0141b
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 252-3714
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-072-429-7

Query Match 65.0%; Score 67; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RPKIWFENRRKPKWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 9
US-08-964-302A-6
Sequence 6, Application US/08964302A
Patent No. 6015787
GENERAL INFORMATION:
APPLICANT: Potter, David A.
APPLICANT: Skolnik, Paul R.
TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,302A
FILING DATE: 04-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 00398/126001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-964-302A-6

Query Match 65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RPKIWFENRRKPKWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 10
US-09-116-294-4
; Sequence 4, Application US/09116294
; Patent No. 6025140
; GENERAL INFORMATION:
; APPLICANT: Langel, Ulo
; APPLICANT: Bartfai, Tamás
; APPLICANT: Pooga, Margus
; APPLICANT: Valkna, Andrus
; APPLICANT: Saar, Kulliki
; APPLICANT: Hallbrink, Mattias
; TITLE OF INVENTION: Conjugated Constructs of Peptides and
; TITLE OF INVENTION: Nucleic Acid Analogs, and Their Transport Across Membranes
; FILE REFERENCE: 4394
; CURRENT APPLICATION NUMBER: US/09/116,294
; CURRENT FILING DATE: 1998-07-16
; EARLIER APPLICATION NUMBER: 60/052,678
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: drosophila
US-09-116-294-4

Query Match 65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKIWFNNRRKPKWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNNRRMKWK 16

RESULT 11
US-08-964-614A-4
; Sequence 4, Application US/08964614A
; Patent No. 6057104
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN
; TITLE OF INVENTION: Rad51 PROTEIN AND DISRUPTION OF PROTEINS THAT ASSOCIATE
; TITLE OF INVENTION: WITH MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,614A
; FILING DATE: 05-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,280
; FILING DATE: 05-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8535-0019-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-614A-4

Query Match 65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKIWFNNRRKPKWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNNRRMKWK 16

RESULT 12
US-08-849-486-1
; Sequence 1, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-849-486-1

Query Match 65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RQPKIWFNNRRKPKWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNNRRMKWK 16

RESULT 13
US-08-849-486-4
; Sequence 4, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:

```
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: FR 95 11714
/ FILING DATE: 05-OCT-1995
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 16 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..16
/ OTHER INFORMATION: /product= "amino acids of the D series"
US-08-849-486-4

Query Match          65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKWFPNRRKPWK 17
Db      1 RQIKWQNRRMKWKK 16

RESULT 14
US-09-208-966-54
/ Sequence 54, Application US/09208966
/ Patent No. 6221355
/ GENERAL INFORMATION:
/ APPLICANT: Dowdy, Steven F.
/ TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
/ FILE REFERENCE: 48881/1742
/ CURRENT APPLICATION NUMBER: US/09/208,966
/ CURRENT FILING DATE: 1998-12-10
/ EARLIER APPLICATION NUMBER: 60/082,402
/ EARLIER FILING DATE: 1998-04-20
/ EARLIER APPLICATION NUMBER: 60/069,012
/ EARLIER FILING DATE: 1997-12-10
/ NUMBER OF SEQ ID NOS: 57
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 54
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: human
US-09-208-966-54

Query Match          65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKWFPNRRKPWK 17
Db      1 RQIKWQNRRMKWKK 16

RESULT 15
US-09-308-935-8
/ Sequence 8, Application US/09308935
/ Patent No. 6268334
/ GENERAL INFORMATION:
/ APPLICANT: La Thangue, Nicholas B
/ APPLICANT: Bandara, Lasantha R
/ TITLE OF INVENTION: Peptide antagonists of DP transcription factors
/ FILE REFERENCE: 620-67
/ CURRENT APPLICATION NUMBER: US/09/308,935
/ CURRENT FILING DATE: 1999-05-27
/ EARLIER APPLICATION NUMBER: PCT/GB97/03506
/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: GB 9626589.7
/ EARLIER FILING DATE: 1996-12-20
/ NUMBER OF SEQ ID NOS: 18
```

```
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-308-935-8

Query Match          65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKWFPNRRKPWK 17
Db      1 RQIKWQNRRMKWKK 16

RESULT 16
US-09-441-416A-6
/ Sequence 6, Application US/09441416A
/ Patent No. 6294518
/ GENERAL INFORMATION:
/ APPLICANT: Potter, David A.
/ APPLICANT: Skolnik, Paul R.
/ TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF
/ TITLE OF INVENTION: CALPAIN
/ FILE REFERENCE: 00398-140001
/ CURRENT APPLICATION NUMBER: US/09/441,416A
/ CURRENT FILING DATE: 1999-11-16
/ PRIOR APPLICATION NUMBER: US 08/964,302
/ PRIOR FILING DATE: 1997-11-04
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-441-416A-6

Query Match          65.0%; Score 67; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKWFPNRRKPWK 17
Db      1 RQIKWQNRRMKWKK 16

RESULT 17
US-09-296-089-33
/ Sequence 33, Application US/09296089
/ Patent No. 6303576
/ GENERAL INFORMATION:
/ APPLICANT: Blaschuk, Orest W.
/ APPLICANT: Byers, Stephen
/ APPLICANT: Gour, Barbara J
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
/ TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
/ FILE REFERENCE: 100086.411
/ CURRENT APPLICATION NUMBER: US/09/296,089
/ CURRENT FILING DATE: 1999-04-21
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 33
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-09-296-089-33

Query Match          65.0%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RQPKWFPNRRKPWK 17
```


US-09-057-363C-47
; Sequence 47, Application US/09057363C
; Patent No. 6551994
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING THE
; INTERACTION BETWEEN ALPHA-CATENIN AND BETA-CATENIN
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,363C
; FILING DATE: 08-Apr-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-057-363C-47

Query Match 55.0%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17
DB 1 RQIKWFOQRNRRMKWK 16

RESULT 23
US-09-043-560B-3
; Sequence 3, Application US/09043560B
; Patent No. 6569833
; GENERAL INFORMATION:
; APPLICANT: Fahraeus, Robin
; APPLICANT: Lane, David P.
; TITLE OF INVENTION: Cyclin Dependent Kinase Binding Compounds
; FILE REFERENCE: CCI-003US
; CURRENT APPLICATION NUMBER: US/09/043,560B
; CURRENT FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-043-560B-3

Query Match 55.0%; Score 67; DB 4; Length 16;

Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 RQPKIWFPPNRRKPWK 17
DB 1 RQIKWFOQRNRRMKWK 16

RESULT 24
US-09-648-400A-29
; Sequence 29, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and into Epithelial Tissues
; FILE REFERENCE: 019801-000210US
; CURRENT APPLICATION NUMBER: US/09/648,400A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Antennapedia
; OTHER INFORMATION: homeodomain, Antennapedia-43-58
US-09-648-400A-29

Query Match 55.0%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17
DB 1 RQIKWFOQRNRRMKWK 16

RESULT 25
US-09-227-652B-4
; Sequence 4, Application US/09227652B
; Patent No. 6610495
; GENERAL INFORMATION:
; APPLICANT: TMV Telethon Institute for Child Health Research
; TITLE OF INVENTION: PEPTIDE DETECTION METHOD
; FILE REFERENCE: 1991209/MRO-PCT
; CURRENT APPLICATION NUMBER: US/09/227,652B
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 60/070989
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Penetratin 16-mer.
US-09-227-652B-4

Query Match 55.0%; Score 67; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17

Db 1 RQIKIWFONRRMVKKK 16

Search completed: May 24, 2004, 17:29:49
Job time : 12.1757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:19:01 ; Search time 31.9324 Seconds
(without alignments)
148.491 Million cell updates/sec

Title: US-09-977-349-2

Perfect score: 103

Sequence: 1 IRQPKWFFNRKPKWK 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	96.1	16	12	US-10-432-291-7
2	99	96.1	16	12	US-10-136-187-22
3	99	96.1	16	14	US-10-136-738-14
4	99	96.1	27	12	US-10-432-291-8
5	93	90.3	16	12	US-10-261-161-28
6	85	82.5	16	9	US-09-981-286A-5
7	82.5	80.1	17	9	US-09-854-204-59
8	75	72.8	16	9	US-09-981-286A-4
9	75	72.8	16	9	US-09-981-286A-4
10	75	72.8	16	14	US-10-261-161-27
11	75	72.8	16	14	US-10-136-738-13
12	69	67.0	17	9	US-09-854-204-22
13	68	66.0	115	9	US-09-925-299-1169
14	68	66.0	115	10	US-09-925-299-1169
15	67	65.0	16	8	US-08-610-220A-9
					Sequence 7, Appli
					Sequence 22, Appli
					Sequence 14, Appli
					Sequence 8, Appli
					Sequence 28, Appli
					Sequence 5, Appli
					Sequence 59, Appli
					Sequence 58, Appli
					Sequence 27, Appli
					Sequence 13, Appli
					Sequence 22, Appli
					Sequence 1169, Ap
					Sequence 9, Appli

16	67	65.0	16	9	US-09-214-371-43	Sequence 43, Appli
17	67	65.0	16	9	US-09-780-070-38	Sequence 38, Appli
18	67	65.0	16	9	US-09-150-623-9	Sequence 9, Appli
19	67	65.0	16	9	US-09-731-023A-10	Sequence 10, Appli
20	67	65.0	16	9	US-09-854-204-1	Sequence 1, Appli
21	67	65.0	16	9	US-09-854-204-56	Sequence 56, Appli
22	67	65.0	16	9	US-09-900-147-8	Sequence 8, Appli
23	67	65.0	16	9	US-09-792-480-29	Sequence 29, Appli
24	67	65.0	16	9	US-09-785-802A-2	Sequence 2, Appli
25	67	65.0	16	9	US-09-785-802A-5	Sequence 5, Appli
26	67	65.0	16	9	US-09-902-432-32	Sequence 32, Appli
27	67	65.0	16	9	US-09-953-031A-10	Sequence 10, Appli
28	67	65.0	16	9	US-09-981-286A-3	Sequence 3, Appli
29	67	65.0	16	10	US-09-982-967A-6	Sequence 6, Appli
30	67	65.0	16	10	US-09-912-414-6	Sequence 6, Appli
31	67	65.0	16	10	US-09-775-052-54	Sequence 54, Appli
32	67	65.0	16	10	US-09-295-189-4	Sequence 4, Appli
33	67	65.0	16	10	US-09-965-876A-1	Sequence 1, Appli
34	67	65.0	16	11	US-09-933-780C-2	Sequence 2, Appli
35	67	65.0	16	12	US-10-286-964-3	Sequence 3, Appli
36	67	65.0	16	12	US-10-432-291-6	Sequence 6, Appli
37	67	65.0	16	12	US-10-357-826A-2	Sequence 2, Appli
38	67	65.0	16	12	US-10-603-409-12	Sequence 12, Appli
39	67	65.0	16	12	US-09-748-063-3	Sequence 3, Appli
40	67	65.0	16	12	US-09-779-791A-3	Sequence 3, Appli
41	67	65.0	16	12	US-10-083-960-29	Sequence 29, Appli
42	67	65.0	16	12	US-10-144-549-1	Sequence 1, Appli
43	67	65.0	16	12	US-10-136-187-21	Sequence 21, Appli
44	67	65.0	16	12	US-10-261-161-1	Sequence 1, Appli
45	67	65.0	16	12	US-10-261-161-26	Sequence 26, Appli
46	67	65.0	16	13	US-10-024-935-12	Sequence 12, Appli
47	67	65.0	16	13	US-10-007-363-3	Sequence 3, Appli
48	67	65.0	16	14	US-10-071-512A-2	Sequence 2, Appli
49	67	65.0	16	14	US-10-239-804-3	Sequence 3, Appli
50	67	65.0	16	14	US-10-209-421-29	Sequence 29, Appli
51	67	65.0	16	14	US-10-077-555-3	Sequence 3, Appli
52	67	65.0	16	14	US-10-229-915-2	Sequence 2, Appli
53	67	65.0	16	14	US-10-185-084-3	Sequence 3, Appli
54	67	65.0	16	14	US-10-252-012-5	Sequence 5, Appli
55	67	65.0	16	14	US-10-075-869-19	Sequence 19, Appli
56	67	65.0	16	14	US-10-013-815-19	Sequence 19, Appli
57	67	65.0	16	14	US-10-136-738-10	Sequence 10, Appli
58	67	65.0	16	14	US-10-210-660-1	Sequence 1, Appli
59	67	65.0	16	14	US-10-210-660-25	Sequence 25, Appli
60	67	65.0	16	14	US-10-156-570A-21	Sequence 21, Appli
61	67	65.0	16	14	US-10-201-394A-14	Sequence 14, Appli
62	67	65.0	16	14	US-10-017-672-11	Sequence 11, Appli
63	67	65.0	16	14	US-10-201-389A-14	Sequence 14, Appli
64	67	65.0	16	14	US-10-161-051-1	Sequence 1, Appli
65	67	65.0	16	14	US-10-358-365-10	Sequence 10, Appli
66	67	65.0	16	14	US-10-061-607A-2	Sequence 2, Appli
67	67	65.0	16	14	US-10-405-339-44	Sequence 44, Appli
68	67	65.0	16	15	US-10-366-493-19	Sequence 19, Appli
69	67	65.0	16	15	US-10-444-662-2	Sequence 2, Appli
70	67	65.0	16	15	US-10-185-593-2	Sequence 2, Appli
71	67	65.0	16	15	US-10-413-160-38	Sequence 38, Appli
72	67	65.0	16	15	US-10-462-138-10	Sequence 10, Appli
73	67	65.0	16	15	US-10-369-226-47	Sequence 47, Appli
74	67	65.0	16	15	US-10-353-678-2	Sequence 2, Appli
75	67	65.0	16	15	US-10-357-529-8	Sequence 8, Appli
76	67	65.0	17	9	US-09-854-204-19	Sequence 19, Appli
77	67	65.0	17	9	US-09-854-204-27	Sequence 27, Appli
78	67	65.0	17	9	US-09-854-204-31	Sequence 31, Appli
79	67	65.0	17	9	US-09-854-204-32	Sequence 32, Appli
80	67	65.0	17	9	US-09-785-802A-3	Sequence 3, Appli
81	67	65.0	17	11	US-09-933-780C-21	Sequence 21, Appli
82	67	65.0	17	12	US-10-602-303-3	Sequence 3, Appli
83	67	65.0	17	12	US-10-421-548-8	Sequence 8, Appli
84	67	65.0	17	13	US-10-007-761-8	Sequence 8, Appli
85	67	65.0	17	14	US-10-209-421-30	Sequence 30, Appli
86	67	65.0	17	14	US-10-229-915-1	Sequence 1, Appli
87	67	65.0	17	14	US-10-210-660-17	Sequence 17, Appli
88	67	65.0	17	14	US-10-210-660-20	Sequence 20, Appli

89 67 65.0 17 14 US-10-210-660-22 Sequence 22, Appl
90 67 65.0 17 14 US-10-210-660-27 Sequence 27, Appl
91 67 65.0 17 15 US-10-372-003A-29 Sequence 29, Appl
92 67 65.0 17 15 US-10-428-280-15 Sequence 15, Appl
93 67 65.0 17 15 US-10-421-503-66 Sequence 66, Appl
94 67 65.0 18 9 US-09-785-802A-14 Sequence 14, Appl
95 67 65.0 19 9 US-09-949-474-7 Sequence 7, Appl
96 67 65.0 19 14 US-10-118-079-45 Sequence 45, Appl
97 67 65.0 19 14 US-10-210-660-23 Sequence 23, Appl
98 67 65.0 19 15 US-10-407-449-20 Sequence 20, Appl
99 67 65.0 20 9 US-09-854-204-63 Sequence 63, Appl
100 67 65.0 20 9 US-09-949-474-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-10-432-291-7
; Sequence 7, Application US/10432291
; Publication No. US20040029281A1
; GENERAL INFORMATION:
; APPLICANT: Centre National de la Recherche Scientifique (CNRS)
; APPLICANT: Joliot, Alain
; APPLICANT: Dupont, Edmond
; APPLICANT: Prochiantz, Alain
; TITLE OF INVENTION: Carrier vectors through an epithelium with tight junctions
; FILE REFERENCE: 45636-5067-US
; CURRENT APPLICATION NUMBER: US/10/432,291
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: PCT/FR01/03631
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: FR 00/14945
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: penetratin sequence for transport vectors
US-10-432-291-7

Query Match 96.1%; Score 99; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.le-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWKX 17
| | | | | | | | | | | | | | | | | |
Db 1 RQPKWFPNRRKPWKX 16

RESULT 2
US-10-136-187-22
; Sequence 22, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 16

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-136-187-22

Query Match 96.1%; Score 99; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.le-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWKX 17
| | | | | | | | | | | | | | | | | |
Db 1 RQPKWFPNRRKPWKX 16

RESULT 3
US-10-136-738-14
; Sequence 14, Application US/10136738
; Publication No. US20030108986A1
; GENERAL INFORMATION:
; APPLICANT: Finn, John
; APPLICANT: MacLachlan, Ian
; APPLICANT: Protiva Biopharmaceutics Inc.
; TITLE OF INVENTION: Autogene Nucleic Acids Encoding a
; TITLE OF INVENTION: Secretable RNA Polymerase
; FILE REFERENCE: 020801-0003100S
; CURRENT APPLICATION NUMBER: US/10/136,738
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,974
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antennapedia homeodomain third helix (residues
; OTHER INFORMATION: 43-58), 3-Pro secretion domain
US-10-136-738-14

Query Match 96.1%; Score 99; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.le-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RQPKWFPNRRKPWKX 17
| | | | | | | | | | | | | | | | | |
Db 1 RQPKWFPNRRKPWKX 16

RESULT 4
US-10-432-291-8
; Sequence 8, Application US/10432291
; Publication No. US20040029281A1
; GENERAL INFORMATION:
; APPLICANT: Centre National de la Recherche Scientifique (CNRS)
; APPLICANT: Joliot, Alain
; APPLICANT: Dupont, Edmond
; APPLICANT: Prochiantz, Alain
; TITLE OF INVENTION: Carrier vectors through an epithelium with tight junctions
; FILE REFERENCE: 45636-5067-US
; CURRENT APPLICATION NUMBER: US/10/432,291
; CURRENT FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: PCT/FR01/03631
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: FR 00/14945
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial

;
; FEATURE:
; OTHER INFORMATION: nuclear export and penetratin sequence for transport vectors
US-10-432-291-8

Query Match 96.1%; Score 99; DB 12; Length 27;

Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ROPKIWFNNRRKPWK 17

Db 12 ROPKIWFNNRRKPWK 27

RESULT 5

US-10-261-161-28
; Sequence 28, Application US/10261161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; TITLE OF INVENTION: Transfer (FRET) Assays For Clostridial Toxins
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-261-161-28

Query Match 90.3%; Score 93; DB 12; Length 16;

Best Local Similarity 93.8%; Pred. No. 6.9e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ROPKIWFNNRRKPWK 17

Db 1 ROPKIWFNNRRKPWK 16

RESULT 6

US-09-981-286A-5
; Sequence 5, Application US/09981286A
; Publication No. US20020192799A1
; GENERAL INFORMATION:
; APPLICANT: Watowich, Stanley J.
; APPLICANT: Weaver, Scott C.
; APPLICANT: Davey, Robert A.
; TITLE OF INVENTION: Drug Discovery Methods
; FILE REFERENCE: 265.00260101
; CURRENT APPLICATION NUMBER: US/09/981,286A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/240,187
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-5

Query Match 82.5%; Score 85; DB 9; Length 16;

Best Local Similarity 87.5%; Pred. No. 8.2e-05;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ROPKIWFNNRRKPWK 17

Db 1 ROPKIWFNNRRKPWK 16

RESULT 7

US-09-854-204-59
; Sequence 59, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CGI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-59

Query Match 80.1%; Score 82.5; DB 9; Length 17;

Best Local Similarity 88.2%; Pred. No. 0.00019;

Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 ROPKIWFNNRRKPWK 17

Db 1 ROPKIWFNNRRKPWK 17

RESULT 8

US-09-854-204-58
; Sequence 58, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CGI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 58
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-58

Query Match 72.8%; Score 75; DB 9; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRKPKWK 17
||| ||||| |||||
Db 1 RQKIWFPPNRKPKWK 16

RESULT 9
US-09-981-286A-4
; Sequence 4, Application US/09981286A
; Publication No. US20020192799A1
; GENERAL INFORMATION:
; APPLICANT: Watowich, Stanley J.
; APPLICANT: Weaver, Scott C.
; TITLE OF INVENTION: Drug Discovery Methods
; FILE REFERENCE: 265.00260101
; CURRENT APPLICATION NUMBER: US/09/981,286A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/240,187
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-4

Query Match 72.8%; Score 75; DB 9; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRKPKWK 17
||| ||||| |||||
Db 1 RQKIWFPPNRKPKWK 16

RESULT 10
US-10-261-161-27
; Sequence 27, Application US/10261161
; Publication No. US2004007270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; FILE REFERENCE: P-Ar 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 16
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-261-161-27

Query Match 72.8%; Score 75; DB 12; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRKPKWK 17
||| ||||| |||||
Db 1 RQKIWFPPNRKPKWK 16

RESULT 11
US-10-136-738-13
; Sequence 13, Application US/10136738
; Publication No. US2003010886A1
; GENERAL INFORMATION:
; APPLICANT: Finn, John
; APPLICANT: MacLachlan, Ian
; TITLE OF INVENTION: Protiva Biotherapeutics Inc.
; TITLE OF INVENTION: Autogene Nucleic Acids Encoding a
; FILE REFERENCE: 020801-000310US
; CURRENT APPLICATION NUMBER: US/10/136,738
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,974
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antennapedia homeodomain third helix (residues
; OTHER INFORMATION: 43-58), Pro50 secretion domain
US-10-136-738-13

Query Match 72.8%; Score 75; DB 14; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRKPKWK 17
||| ||||| |||||
Db 1 RQKIWFPPNRKPKWK 16

RESULT 12
US-09-854-204-22
; Sequence 22, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750

; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: MOD RES
; LOCATION: (1)-
; OTHER INFORMATION: bala
; NAME/KEY: MOD RES
; LOCATION: (17)
; OTHER INFORMATION: AMIDATION
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-22

Query Match 67.0%; Score 69; DB 9; Length 17;
Best Local Similarity 75.0%; Pred. No. 0.012;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPWK 17
DB 2 RQKIWFNRRKPWK 17

RESULT 13
US-09-925-299-1169
; Sequence 1169, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1169
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1169

Query Match 66.0%; Score 68; DB 9; Length 115;
Best Local Similarity 75.0%; Pred. No. 0.089;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPWK 17
DB 64 RQKIWFNRRKPWK 79

RESULT 14
US-09-925-299-1169
; Sequence 1169, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1169
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1169

Query Match 66.0%; Score 68; DB 10; Length 115;
Best Local Similarity 75.0%; Pred. No. 0.089;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPWK 17
DB 64 RQKIWFNRRKPWK 79

RESULT 15
US-08-610-220A-9
; Sequence 9, Application US/08610220A
; Publication No. US20030099638A1
; GENERAL INFORMATION:
; APPLICANT: Troy, Carol M.
; TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/610,220A
; FILING DATE: MAR-04-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 48332/JPW/JML
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-610-220A-9

Query Match 65.0%; Score 67; DB 8; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFNRRKPWK 17
DB 1 RQKIWFNRRKPWK 16

RESULT 16
US-09-214-371-43
; Sequence 43, Application US/09214371B
; Patent No. US20010018511A1
; GENERAL INFORMATION:
; APPLICANT: Lane, David
; APPLICANT: Bottger, Volker

APPLICANT: Bottger, Angelica
APPLICANT: Picksley, Stephen
APPLICANT: Chene, Patrick
APPLICANT: Hochkeppel, Heinz-Kurt
APPLICANT: Garcia-Echeverria, Carlos
APPLICANT: Furet, Pascal
TITLE OF INVENTION: Inhibitors of the Interaction of P53 and MDM2
FILE REFERENCE: 4-20937/A/PCT
CURRENT APPLICATION NUMBER: US/09/214,371B
CURRENT FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: PCT/EP97/03549
PRIOR FILING DATE: 1997-07-04
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide
US-09-214-371-43

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRRKPWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 17

US-09-780-070-38
Sequence 38, Application US/09780070
Patent No. US20020009752A1
GENERAL INFORMATION:
APPLICANT: Burke, James
APPLICANT: Strittmater, Warren
APPLICANT: Nagai, Yoshitaka
TITLE OF INVENTION: COMPOUNDS THAT SELECTIVELY BIND TO EXPANDED POLYGLUTAMINE REPEAT
FILE REFERENCE: 5405.242
CURRENT APPLICATION NUMBER: US/09/780,070
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/189,781
PRIOR FILING DATE: 2000-03-16
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 16
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-780-070-38

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRRKPWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 18

US-09-150-623-9
Sequence 9, Application US/09150623
Patent No. US20020044931A1
GENERAL INFORMATION:
APPLICANT: Troy, Carol M.
TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL
DEATH AND USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,623
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/610,220
FILING DATE: MAR-04-1996
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 48332/JPW/JML
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-150-623-9

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRRKPWK 17
DB 1 RQKIWFQNRMRKWK 16

RESULT 19

US-09-731-023A-10
Sequence 10, Application US/09731023A
Patent No. US20020077283A1
GENERAL INFORMATION:
APPLICANT: Sessa, William
TITLE OF INVENTION: Caveolin Peptides and Their Use as Therapeutics
FILE REFERENCE: 44574-5076-US
CURRENT APPLICATION NUMBER: US/09/731,023A
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/231,327
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 16
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(16)
OTHER INFORMATION: Homeodomain, internalization sequence
US-09-731-023A-10

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQKIWFPPNRRKPWK 17
DB 1 RQKIWFQNRMRKWK 16

Db 1 RQIKIWFONRRMKNKK 16

RESULT 20

```

US-09-854-204-1
; Sequence 1, Application US/09854204
; Patent No. US2002098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE OF INVENTION: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-854-204-1

```

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels

Qy 2 RQPKIWFNRRKPKWK 17
|||
db 1 RQIKIWFONRRMKWK 16

RESULT 21

US-09-854-204-56
 ; Sequence 56, Application US/09854204
 ; Patent NO. US20020098236A1
 ; GENERAL INFORMATION:
 APPLICANT: Fischer, Peter Martin
 APPLICANT: Zhelev, Nikolai
 TITLE OF INVENTION: Transport Vectors
 FILE REFERENCE: CCI-010
 CURRENT APPLICATION NUMBER: US/09/854,204
 CURRENT FILING DATE: 2001-05-11
 PRIOR APPLICATION NUMBER: 09/438,460
 PRIOR FILING DATE: 1999-11-12
 PRIOR APPLICATION NUMBER: BG 9825000.4
 PRIOR FILING DATE: 1998-11-13
 PRIOR APPLICATION NUMBER: BG 9825001.2
 PRIOR FILING DATE: 1998-11-13
 PRIOR APPLICATION NUMBER: BG 9802525.6
 PRIOR FILING DATE: 1999-02-04
 PRIOR APPLICATION NUMBER: BG 9902522.3
 PRIOR FILING DATE: 1999-02-04
 PRIOR APPLICATION NUMBER: BG 9914578.1
 PRIOR FILING DATE: 1999-06-22
 PRIOR APPLICATION NUMBER: PCT/GB99/03750
 PRIOR FILING DATE: 1999-11-11
 NUMBER OF SEQ ID NOS: 66
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 56

```

; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (12)
; OTHER INFORMATION: Xaa is norleucine ornithine
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-56

```

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels

Qy 2 RQPKIWFENRRKPWK 17
db 1 RQIKIWFONRRXKWK 16

RESULT 22

```

US-09-900-147-8
; Sequence 8, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-900-147-8

```

Query Match	65.0%	Score 67;	DB 9;	Length 16;
Best Local Similarity	75.0%;	Pred. No. 0.021;		
Matches 12;	Conservative	0;	Mismatches 4;	Indels 0;

Qy 2 RQPKIWFPNRRKPWK 17
||| ||| ||| |||
Db 1 ROIKWFOERNMKWK 16

RESULT 23

```

US-09-792-480-29
; Sequence 29, Application US/09792480
; Patent No. US20020127198A1
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Sista, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: CellGate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000230US
; CURRENT APPLICATION NUMBER: US/09/792,480
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 09/648,400
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 57

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Antennapedia
; OTHER INFORMATION: homeodomain, Antennapedia-43-58
US-09-792-480-29

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNRMRMKWK 16

RESULT 24
US-09-785-802A-2
; Sequence 2, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Membrane translocation sequence from Penetratin
US-09-785-802A-2

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNRMRMKWK 16

RESULT 25
US-09-785-802A-5
; Sequence 5, Application US/09785802A
; Patent No. US20020151004A1
; GENERAL INFORMATION:
; APPLICANT: Craig, Roger
; TITLE OF INVENTION: DELIVERY VEHICLES AND METHODS FOR USING THE SAME
; FILE REFERENCE: 11067/2035
; CURRENT APPLICATION NUMBER: US/09/785,802A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/748,06
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/748,789
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-785-802A-5

Query Match 65.0%; Score 67; DB 9; Length 16;
Best Local Similarity 75.0%; Pred. No. 0.021;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RQPKIWFPPNRRKPWK 17
||| ||| ||| |||
Db 1 RQIKIWFQNRMRMKWK 16

Search completed: May 24, 2004, 17:28:44
Job time : 31.9324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:13:28 ; Search time 53.5135 Seconds
(without alignments)
105.599 Million cell updates/sec

Title: US-09-977-349-3
Perfect score: 118
Sequence: 1 CSSCRQIKIWFQNRMRMKWK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	101	85.6	17	3	ADE14787
2	101	85.6	17	3	ADE14777
3	101	85.6	17	5	ABG31156
4	101	85.6	17	6	ABR43083
5	101	85.6	18	3	AAV93956
6	101	85.6	19	2	AAV22240
7	101	85.6	19	4	ABR48441
8	99	83.9	233	7	ABE56916
9	99	83.9	233	7	ABE56920
10	99	83.9	270	7	ABE56922
11	99	83.9	270	7	ABE56918
12	98	83.1	20	5	AAU98543
13	98	83.1	20	5	AAU63051
14	96	81.4	60	2	AAV42291
15	96	81.4	60	2	AAV04364
16	96	81.4	60	2	RAY27403
17	96	81.4	60	5	ABR4470
18	96	81.4	61	4	AAE00811
19	96	81.4	64	6	AAE32060
20	96	81.4	128	6	ABR99729
21	96	81.4	217	4	AAE10922
22	96	81.4	217	5	ABG96337
23	96	81.4	230	3	AAV58610
24	96	81.4	295	6	AAE32047
25	96	81.4	378	4	ABB72035

26	96	81.4	417	4	ABB57755
27	93	78.8	24	3	AAI17276
28	93	78.8	24	5	ABB73354
29	93	78.8	30	4	AAI17276
30	92	78.0	16	2	AAW45974
31	92	78.0	16	2	AAW33407
32	92	78.0	16	2	AAW33410
33	92	78.0	16	2	AAW82958
34	92	78.0	16	2	AAW56397
35	92	78.0	16	2	AAW71270
36	92	78.0	16	2	AAW71316
37	92	78.0	16	2	AAW30508
38	92	78.0	16	2	AAW91046
39	92	78.0	16	2	AAV52102
40	92	78.0	16	2	AAV00859
41	92	78.0	16	2	AAV13509
42	92	78.0	16	3	AAV87920
43	92	78.0	16	3	AAE27050
44	92	78.0	16	3	AAV3667
45	92	78.0	16	3	AAV67966
46	92	78.0	16	3	AAV93551
47	92	78.0	16	3	AAV55818
48	92	78.0	16	3	AAV71008
49	92	78.0	16	3	AAV51212
50	92	78.0	16	3	AAV51167
51	92	78.0	16	3	AAI10343
52	92	78.0	16	3	AAI19251
53	92	78.0	16	3	AAV93178
54	92	78.0	16	3	AAV35694
55	92	78.0	16	3	AAE22025
56	92	78.0	16	3	AAE29423
57	92	78.0	16	3	AAE03927
58	92	78.0	16	3	AAV93954
59	92	78.0	16	3	AAE29574
60	92	78.0	16	3	ADE14785
61	92	78.0	16	3	ADE14761
62	92	78.0	16	4	AAV73091
63	92	78.0	16	4	AAE60004
64	92	78.0	16	4	AAE70753
65	92	78.0	16	4	AAE02974
66	92	78.0	16	4	AAE60671
67	92	78.0	16	4	AAU06064
68	92	78.0	16	4	AAU49914
69	92	78.0	16	4	AAE66396
70	92	78.0	16	4	AAU00813
71	92	78.0	16	4	AAE12205
72	92	78.0	16	5	ABB78030
73	92	78.0	16	5	ABG78985
74	92	78.0	16	5	ABG78985
75	92	78.0	16	5	AAE15616
76	92	78.0	16	5	AAE15616
77	92	78.0	16	5	AAE15781
78	92	78.0	16	5	AAU78345
79	92	78.0	16	5	ABG68406
80	92	78.0	16	5	ABE83153
81	92	78.0	16	5	ABP53809
82	92	78.0	16	5	AAE25451
83	92	78.0	16	5	ABG75500
84	92	78.0	16	5	ABG60447
85	92	78.0	16	5	ABG30763
86	92	78.0	16	5	AAU78912
87	92	78.0	16	5	ABE81174
88	92	78.0	16	5	AAU77234
89	92	78.0	16	5	ABG5801
90	92	78.0	16	5	ABP54754
91	92	78.0	16	5	AAE47571
92	92	78.0	16	5	AAE48196
93	92	78.0	16	5	AAE23684
94	92	78.0	16	6	ABP72298
95	92	78.0	16	6	ABM32768
96	92	78.0	16	6	ABG73390
97	92	78.0	16	6	ABE82532
98	92	78.0	16	6	ABB99502

Abb57755 Drosophil
Aab17276 p16-mimet
Abb73354 Exemplary
Aab31443 Peptide d
Aaw45974 Cysteine
Aaw33407 Peptide 4
Aaw33410 D-form pe
Aaw82958 Oestrogen
Aaw56397 Preferred
Aaw71270 Antennape
Aaw71316 Antennape
Aaw30508 Drosophil
Aaw91046 Internali
Aay52102 Peptide f
Aay00859 Peptide p
Aay13509 Signal se
Aay87920 Drosophil
Aab27060 Beta-cate
Aay93667 Peptide w
Aay93667 Peptide w
Aay93551 Amino aci
Aay55818 Signal se
Aay71008 Drosophil
Aay51212 Antennape
Aay51167 Drosophil
Aab10343 Peptide A
Aab19251 Fragment
Aay93178 Protegrin
Aab35694 Peptide a
Aab22025 Membrane
Aab29423 ANTP pept
Aab03927 Internali
Aay93954 Peptide u
Aay29574 Antennape
Ade14785 Carrier m
Ade14761 Drosophil
Aab73091 Rheumatol
Aab60004 Internali
Aab70753 Cell memb
Aae02974 Protein t
Aab60671 Antennape
Aau06064 Drosophil
Aau49914 Hif-1alph
Aab66396 Antennape
Aau00813 Fruit fly
Aae12205 Membrane
Abb78030 Peptide d
Abb78985 Cell pene
Abb78214 Amino aci
Aae15616 Drosophil
Aae15781 Antennape
Aau78345 Antennape
Abg68406 Transloca
Abb83153 Transduct
Abb53809 Penetrati
Aae25451 Drosophil
Abg75500 Fruit fly
Abg60447 Selective
Abg30763 E2F-1/cyc
Aau78912 Antennape
Abb81174 Antennape
Aau77234 Drosophil
Abb95801 Cell pene
Abp54754 Antennape
Aam47571 Penetrati
Aam48196 Drosophil
Aae23684 Fluoresce
Aap72298 ANTENNAP
Abm32768 DGI-3 spe
Aae35385 Antennape
Abg73390 Penetrati
Abb82532 Drosophil
Abb99502 Sequence

99 92 78.0 16 6 AA019913 Antennapedia homeodomain of Drosophila melanogaster. The drug and carrier moieties associated with this sequence in the exemplification of the invention are featured.

100 92 78.0 16 6 ABG76121 Fruitfly

ALIGNMENTS

RESULT 1
ADE14787
ID ADE14787 standard; peptide; 17 AA.

AC ADE14787;

XX

XX

DT 29-JAN-2004 (first entry)

XX

DE

XX

KW drug delivery system; carrier group; homeobox peptide; drug metabolism;

KW drug biodistribution; drug excretion; drug half life; drug solubility;

KW drug side-effect; drug resistance; penatrin; antennapedia homeodomain;

XX fruit fly; mutant; mutein.

OS Synthetic.

OS Drosophila melanogaster.

XX

PH Key

Modified-site

1

/note= "Optionally C-terminal H; Optionally bound to 2'-

-(succinimidopropionyl)-paclitaxel; Optionally bound to 4

bound to 4'-(succinimidopropionyl)-epipodophyllotoxin;

Optionally bound to 4'-Dimethyl-4-(acetyl)-

epipodophyllotoxin; Optionally bound to 10-O-

(succinimidopropionyl)-camptothecin; Optionally bound

to 4'Methoxy-4-(4'-aminoanilino-

(succinimidopropionyl)-epipodophyllotoxin"

2. 17

/note= "Optionally these residues are all D-form

residues"

17

Modified-site

/note= "N-terminal amide"

XX

PN W0200001417-A1.

XX

PD 13-JAN-2000.

XX

PF 22-JUN-1999; 99WO-GB001957.

XX

PR 03-JUL-1998; 98GB-00014527.

XX

PA (CYCL-) CYCLACEL LTD.

XX

PI Fischer PM, Wang S;

XX

DR WPI; 2000-160729/14.

XX

PT Delivery systems with improved delivery into target cells, half-life,

PT solubility in biological fluids, enhanced metabolism, biodistribution and

PT excretion, and minimal toxic side-effects.

XX

PS Example 6; SEQ ID NO 27; 114pp; English.

XX

CC This invention relates to novel drug delivery systems comprising a drug

CC group linked to a carrier group comprising a homeobox peptide, its

CC fragment or derivative. The invention may provide a method of improved

CC delivery into target cells with enhanced metabolism, biodistribution and

CC excretion of a drug. The drug may be therapeutically active in both

CC intact and dissociated states. The invention may improve half-life of a

CC drug in human or animal bodies, improve solubility in biological fluids,

CC minimise known toxic or non-desirable side-effects, enhance onset of

CC action of a desired therapeutic effect, provide alternative routes for

CC administration of drug and decrease incidence of drug resistance. The

CC present sequence is the amino acid sequence of a carrier peptide derived

CC from penatrin, a peptide derived from the antennapedia homeodomain of

CC Drosophila melanogaster. The drug and carrier moieties associated with

CC this sequence in the exemplification of the invention are featured.

XX

SQ Sequence 17 AA;

Query Match 95.6%; Score 101; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CRQIKWFOQRMRMKWK 20

|||||

Db 1 CRQIKWFOQRMRMKWK 17

RESULT 2

ADE14777

ID ADE14777 standard; peptide; 17 AA.

XX

AC ADE14777;

XX

DT 29-JAN-2004 (first entry)

XX

DE

XX

KW drug delivery system; carrier group; homeobox peptide; drug metabolism;

KW drug biodistribution; drug excretion; drug half life; drug solubility;

KW drug side-effect; drug resistance; penatrin; antennapedia homeodomain;

XX fruit fly; mutant; mutein.

OS Synthetic.

OS Drosophila melanogaster.

XX

PH Key

Modified-site

1

/note= "Optionally N-terminal H; Optionally labelled with

N alpha trityl (Trt); Optionally bound to N-(3-(3-

succinimido)benzoyl)-doxorubicin; Optionally bound to 2'-

(succinimidopropionyl)-paclitaxel; Optionally bound to 2'-

2'-Methoxyacetyl-7-(succinimidopropionyl)-paclitaxel;

Optionally bound to 7-(succinimidopropionyl)-paclitaxel;

Optionally bound to 4'-(succinimidopropionyl)-

epipodophyllotoxin; Optionally bound to 4'-Chloroacetyl-4

-(succinimidopropionyl)-epipodophyllotoxin; Optionally

bound to 4'-Dimethyl-4-(succinimidopropionyl)-

epipodophyllotoxin; Optionally bound to G2-

(succinimidopropionyl)-etoposide; Optionally bound to G3

4'-(succinimidopropionyl)-etoposide; Optionally bound to

O-(succinimidopropionyl)-roscovitine; Optionally bound to

to O-(succinimidopropionyl)-bohemine; Optionally bound

to 4-(acetyl)-podophyllotoxin; Optionally bound to 2'-(3-

(3-succinimido)propionyl)-paclitaxel; Optionally bound

to O-(3-(3-succinimido)propionyl)-bohemine"

2

/note= "Optionally labelled with 2,2,5,7,8-

pentamethylchroman-6-sulfonyl (Pmc)"

3

Modified-site

/note= "Optionally labelled with N alpha trityl (Trt)"

5

Modified-site

/note= "Optionally labelled with t-Butyloxycarbonyl

(Boc)"

9

Modified-site

/note= "Optionally labelled with N alpha trityl (Trt)"

10

Modified-site

/note= "Optionally labelled with N alpha trityl (Trt)"

11

Modified-site

/note= "Optionally labelled with 2,2,5,7,8-

pentamethylchroman-6-sulfonyl (Pmc)"

12

Modified-site

/note= "Optionally labelled with 2,2,5,7,8-

pentamethylchroman-6-sulfonyl (Pmc)"

14

Modified-site

FT FT Modified-site /note= "Optionally labelled with t-Butyloxy carbonyl
 FT (Boc)"
 FT 16
 FT /note= "Optionally labelled with t-Butyloxy carbonyl
 FT (Boc)"
 FT 17
 FT /note= "Optionally C-terminal OH; Optionally bound to
 FT resin; Optionally labelled with t-Butyloxy carbonyl (Boc)"
 FT 18
 PN WO200001417-A1.
 XX 13-JAN-2000.
 PD 22-JUN-1999; 99WO-GB001957.
 XX 03-JUL-1998; 98GB-00014527.
 XX (CYCL-) CYCLACEL LTD.
 PA Fischer PM, Wang S;
 PI WPI; 2000-160729/14.
 DR Delivery systems with improved delivery into target cells, half-life,
 XX solubility in biological fluids, enhanced metabolism, biodistribution and
 PT excretion, and minimal toxic side-effects.
 PT Example 2; SEQ ID NO 17; 114pp; English.
 PS This invention relates to novel drug delivery systems comprising a drug
 XX group linked to a carrier group comprising a homeobox peptide, its
 CC fragment or derivative. The invention may provide a method of improved
 CC delivery into target cells with enhanced metabolism, biodistribution and
 CC excretion of a drug. The drug may be therapeutically active in both
 CC intact and dissociated states. The invention may improve half-life of a
 CC drug in human or animal bodies, improve solubility in biological fluids,
 CC minimize known toxic or non-desirable side-effects, enhance onset of
 CC action of a desired therapeutic effect, provide alternative routes for
 CC administration of drug and decrease incidence of drug resistance. The
 CC present sequence is the amino acid sequence of a carrier peptide derived
 CC from penatrin, a peptide derived from the antennapedia homeodomain of
 CC Drosophila melanogaster. The drug and carrier moieties associated with
 CC this sequence in the exemplification of the invention are featured.
 XX SQ Sequence 17 AA;
 Query Match 85.6%; Score 101; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CROIKIWFQNRMRKWK 20
 Db 1 CROIKIWFQNRMRKWK 17
 RESULT 3
 ABG31156
 ID ABG31156 standard; peptide; 17 AA.
 XX ABG31156;
 AC 06-AUG-2003 (revised)
 XX 05-NOV-2002 (first entry)
 DT Drosophila antennapedia homeodomain-derived carrier peptide.
 DE Delta protein kinase C; deltaPKC; V1 domain; vasotropic;
 XX cerebrotrophic; deltaV1-1; deltaV1-2; RACK; pseudo-delta RACK;
 KW pseudo-delta receptor for activated C-kinase; deltaV1-5; PKC;
 KW protein kinase C; signal transduction; cell growth; gene expression;
 KW ion channel activity; translocation; hypoxia; stroke; ischaemic damage;
 KW creatine kinase; homeodomain; carrier peptide.
 XX

OS Drosophila sp.
 OS Synthetic.
 XX WO200257413-A2.
 PN 25-JUL-2002.
 XX 09-NOV-2001; 2001WO-US047556.
 PF 18-JAN-2001; 2001US-0262060P.
 PR (STED) UNIV LELAND STANFORD JUNIOR.
 XX Mochly-Rosen D;
 PI WPI; 2002-599715/64.
 DR New delta protein kinase C peptide for reducing or enhancing damage to
 XX cells or tissues exposed to ischemic or hypoxic event caused by stroke,
 PT or for protecting tissue from damage due to ischemia.
 PT Example 1; Page 53; 65pp; English.
 PS The invention discloses peptides comprising deltaV1-1, deltaV1-2, pseudo-
 CC delta receptors for activated C-kinase (RACK), deltaV1-5 or their
 CC derivatives or fragments. Protein kinase C (PKC) is a key enzyme in
 CC signal transduction involved in a variety of cellular functions including
 CC cell growth, regulation of gene expression and ion channel activity. The
 CC localisation of different PKC isozymes to different areas of the cell in
 CC turn appears due to binding of the activated isozymes to the specific
 CC anchoring molecules (RACKs). Peptides that mimic either the PKC-binding
 CC site on RACKs or the RACK-binding site on PKC are isozyme specific
 CC translocation inhibitors of PKC. The disclosed peptides are useful in
 CC activating or inhibiting translocation or function of deltaPKC. The
 CC deltaPKC agonists or antagonists are useful in reducing, enhancing or
 CC protecting against damage to cells or tissues due to ischaemic or hypoxic
 CC event caused by stroke. Acute administration of the peptides, conjugated
 CC to a carrier peptide or a Tat-derived peptide, protected hearts against
 CC ischaemic damage as shown by decreased release of creatine kinase. The
 CC data indicate that in an intact heart, inhibition of deltaPKC conferred
 CC greater than 50% protection against ischaemic damage. The sequence
 CC presented is the Drosophila antennapedia homeodomain-derived carrier
 CC peptide, which was used to facilitate transport of the rat delta protein
 CC kinase C (PKC) agonists or antagonists across the cell membrane. (Updated
 CC on 06-AUG-2003 to correct OS field.)
 XX SQ Sequence 17 AA;
 Query Match 85.6%; Score 101; DB 5; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CROIKIWFQNRMRKWK 20
 Db 1 CROIKIWFQNRMRKWK 17
 RESULT 4
 ABR43083
 ID ABR43083 standard; peptide; 17 AA.
 XX ABR43083;
 AC 03-JUL-2003 (first entry)
 XX Drosophila antennapedia homeodomain-derived carrier peptide SEQ ID NO:12.
 DE Human; telomerase inhibitory peptide; TEIPP; hTERT; inhibition; cancer;
 KW telomerase reverse transcriptase; cytostatic; telomerase inhibitor;
 KW proliferative disorder; inducing cell death; cell proliferation;
 KW cell growth inhibition.
 XX Drosophila sp.
 OS

OS Synthetic.
 XX WO2003024997-A1.
 PN
 XX
 PD 27-MAR-2003.
 XX
 PF 13-SEP-2002; 2002WO-AU001263.
 XX
 PR 14-SEP-2001; 2001AU-00007697.
 XX
 PA (LIUJ/) LIU J.
 PA (LIHH/) LI H.
 XX
 PI Liu J, Li H;
 XX
 DR WPI; 2003-354586/33.
 XX
 XX New telomerase inhibitory peptides useful for treating or preventing
 PT cancer and other proliferative disorder, for inducing cell death or cell
 PT proliferation, and for inhibiting cell growth.
 XX
 PS Example 2; Page 23; 50pp; English.
 XX
 CC The present invention describes a telomerase inhibitory peptide (TEIPP).
 CC TEIPP is derived from human telomerase reverse transcriptase (hTERT).
 CC TEIPP has cytostatic activity, and can be used as a telomerase inhibitor.
 CC The TEIPP peptide is useful for treating or preventing cancer and other
 CC proliferative disorders, for inducing cell death or cell proliferation,
 CC and for inhibiting cell growth. The present sequence represents a
 CC Drosophila antennapedia homeodomain-derived carrier peptide, which is
 CC used in an example from the present invention
 XX
 SQ Sequence 17 AA;
 XX
 Query Match 85.6%; Score 101; DB 6; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CRQIKIWFONRRMKWK 20
 DB 1 CRQIKIWFONRRMKWK 17
 XX
 RESULT 5
 AAY93956
 ID AAY93956 standard; peptide; 18 AA.
 AC AAY93956;
 XX
 DT 03-OCT-2000 (first entry)
 XX
 DE Peptide used to construct inhibitory pseudopeptide compounds.
 XX
 KW Pseudopeptide compound; inhibitor; tyrosine kinase;
 XX growth factor receptor binding protein; Grb2; proliferative disorder;
 KW cancer; metastasis; antennapedia protein.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Disulfide-bond 1..2
 FT
 XX WO200039153-A1.
 PN
 XX
 PD 06-JUL-2000.
 XX
 XX 24-DEC-1999; 99WO-FR003289.
 PF
 XX
 PR 24-DEC-1998; 98FR-00016459.
 XX
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA
 XX

PI Garbay C, Liu W, Vidal M, Roques BP;
 XX WPI; 2000-475683/41.
 XX
 DR Pseudopeptides, useful for the treatment of proliferative process
 XX disorders, cancers, and metastases, inhibit paths activated by proteins
 PT with a tyrosine kinase activity.
 PT
 XX Claim 1; Page 32; 43pp; French.
 PS
 XX The specification describes pseudopeptide compounds. These compounds are
 CC inhibitors of routes activated by proteins having a tyrosine kinase
 CC activity. The compounds also have high affinity for a growth factor
 CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
 CC the treatment of disorders of proliferative processes, cancers, and
 CC metastases. The present sequence is derived from the antennapedia
 CC protein, and is a fragment of the pseudopeptide compounds of the
 CC invention
 XX
 SQ Sequence 18 AA;
 XX
 Query Match 85.6%; Score 101; DB 3; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CRQIKIWFONRRMKWK 20
 DB 2 CRQIKIWFONRRMKWK 18
 XX
 RESULT 6
 AAY22240
 ID AAY22240 standard; protein; 19 AA.
 AC AAY22240;
 XX
 DT 20-SEP-1999 (first entry)
 XX
 DE Drosophila penetratin motif.
 XX
 KW Peptide detection; pBLOCK-1; biological interaction; bacterial protein;
 KW modulator identification; oncoprotein; leukaemia; solid tumour; therapy;
 KW penetratin motif.
 XX
 OS Drosophila sp.
 XX
 PN WO9935282-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 08-JAN-1999; 99WO-AU000018.
 XX
 PR 09-JAN-1998; 98US-0070989P.
 XX
 PA (TWNT-) TVM TELETHON INST CHILD HEALTH RES.
 XX
 PI Watt PM, Kees UR;
 XX
 DR WPI; 1999-430405/36.
 XX
 PT Detecting peptides that modulate biological interaction, for identifying
 PT antibacterial peptides and antitumor agents.
 XX
 PS Disclosure; Page 17; 77pp; English.
 XX
 CC This sequence represents the Drosophila penetratin motif. The invention
 CC relates to a method for identifying a (oligo)peptide or polypeptide (I)
 CC able to modulate a biological interaction (BI) in a host cell by
 CC producing a peptide library in transformed cells that include: (i) a
 CC nucleic acid (N1) comprising, or encoding, a reporter molecule, the
 CC expression of which is under control of BI; and (ii) a second sequence
 CC (N2) that encodes (I), under control of a promoter. The cells are
 CC cultured to express N2, then any cells in which expression of the

CC reporter is altered are selected. The method is used to identify agents
 CC that modulate BI, i.e. potential diagnostic, prophylactic or therapeutic
 CC agents, particularly where the interactions involve proteins and/or
 CC nucleic acids. Particularly (I) interfere with interactions involving
 CC oncoproteins, for treatment of leukaemia and/or solid tumours, or
 CC bacterial proteins. This method allows rapid, large-scale screening of
 CC libraries for (I) and the effect of library compounds on any selected
 CC interaction can be determined. (I) should be selective, with few side
 CC effects
 CC
 CC SQ Sequence 19 AA;

Query Match 85.6%; Score 101; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CROIKIWFQNRMKKK 20
 |||||
 Db 1 CROIKIWFQNRMKKK 17

RESULT 7
 AAB48441
 ID AAB48441 standard; peptide; 19 AA.
 AC AAB48441;
 DT 02-MAR-2001 (first entry)
 DE Drosophila penetratin targeting sequence.
 KW Drosophila; penetratin targeting sequence; antibacterial; antifungal;
 KW antiviral; allergy; autoimmune disease; antibiotic;
 KW antigenicity modulator; immunogenicity modulator.
 XX Drosophila sp.

Key Location/Qualifiers
 FH Misc-difference 18 /note= "Xaa= any amino acid, any number of residues"
 FT WO200068373-A1.
 XX 16-NOV-2000.
 XX 05-MAY-2000; 2000WO-AU000414.
 XX 05-MAY-1999; 99US-0132711P.
 XX (TWMT-) TVM TELETHON INST CHILD HEALTH RES.

PA Watt PM, Thomas WR;
 PI WPI; 2001-016093/02.
 DR Identifying a modulator or mediator of antigenicity and/or immunogenicity
 PT useful as vaccines or agents for immunotherapy of allergic or autoimmune
 PT disease, by using biodiverse gene fragment libraries.
 XX Disclosure; Page 49; 96pp; English.

XX The present sequence may be fused to a protein in order to assist
 CC penetration or uptake of the protein. The sequence is given in a
 CC specification relating to a method for identifying a modulator or
 CC mediator of a biological activity such as antigenicity and/or
 CC immunogenicity. The method comprises producing a gene fragment expression
 CC library from defined nucleotide sequence fragments and assaying the
 CC library for an amino acid sequence for a biological activity that is
 CC different from any activity the amino acid sequence has in its native
 CC environment. The identified modulators are useful as vaccines or agents
 CC for immunotherapy of allergic or autoimmune disease, antibiotic or
 CC inhibitory agents

SQ Sequence 19 AA;

Query Match 85.6%; Score 101; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CROIKIWFQNRMKKK 20
 |||||
 Db 1 CROIKIWFQNRMKKK 17

RESULT 8
 ADE56916
 ID ADE56916 standard; protein; 233 AA.
 XX AC ADE56916;
 XX DT 29-JAN-2004 (first entry)
 XX DE Rat Protein P52949, SEQ ID NO 2771.
 XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX OS Rattus norvegicus.

XX WO2003016475-A2.
 XX 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
 WPI; 2003-268312/26.
 DR GENBANK; P52949.

PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 233 AA;

Query Match 83.9%; Score 99; DB 7; Length 233;
 Best Local Similarity 90.0%; Pred. No. 5.8e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKIWFQNRMRKWK 20
 Db 212 CLSERQIKIWFQNRMRKWK 231

RESULT 9
 ADE56920
 ID ADE56920 standard; protein; 233 AA.

XX AC ADE56920;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P52949, SEQ ID NO 2775.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; P52949.

XX PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 233 AA;

Query Match 83.9%; Score 99; DB 7; Length 233;
 Best Local Similarity 90.0%; Pred. No. 5.8e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKIWFQNRMRKWK 20
 Db 212 CLSERQIKIWFQNRMRKWK 231

RESULT 10
 ADE56922
 ID ADE56922 standard; protein; 270 AA.

XX AC ADE56922;

XX DT 29-JAN-2004 (first entry)

XX DE Human Protein P20719, SEQ ID NO 2777.

XX KW Human; pain; neuronal tissue; gene therapy;
 XX spinal segmental nerve injury; chronic constriction injury; CCI;
 XX spared nerve injury; SNI; Chung.

XX OS Homo sapiens.

XX PN W02003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-268312/26.

XX DR GENBANK; P20719.

XX PT New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (CCI) in an animal (e.g. gene
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 270 AA;

Query Match 83.9%; Score 99; DB 7; Length 270;
 Best Local Similarity 90.0%; Pred. No. 6.6e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20
 Db 233 CLSERQIKIWFQNRMKWK 252

RESULT 11
 ADE56918
 ID ADE56918 standard; protein; 270 AA.
 XX
 AC ADE56918;

DT 29-JAN-2004 (first entry)

DE Human Protein P20719, SEQ ID NO 2773.

KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

FR 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; P20719.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 270 AA;

Query Match 83.9%; Score 99; DB 7; Length 270;
 Best Local Similarity 90.0%; Pred. No. 6.6e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20

Db 233 CLSERQIKIWFQNRMKWK 252

RESULT 12

AAU98543

ID AAU98543 standard; peptide; 20 AA.

XX AC AAU98543;

XX 23-AUG-2002 (first entry)

DE Peptide linked oligomer compound related peptide #2.

KW Peptide linked oligomeric compound;

KW phosphorothioate 2'-O-MOE gapmer oligonucleotide.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 2

FT /label= OTHER

FT /note= "OTHER= aminobutyric acid"

XX WO200220544-A1.

XX 14-MAR-2002.

XX 07-SEP-2001; 2001WO-US028083.

XX 08-SEP-2000; 2000US-00658517.

XX (ISIS-) ISIS PHARM INC.

XX Manoharan M, Guzaev AP;

XX WPI; 2002-489670/52.

XX Preparing peptide linked oligomeric compound useful for diagnostics,
 CC therapeutics and as research reagents and kits by employing equimolar
 CC amounts functionalized oligomeric compounds and peptide reagents.

XX Example 20; Page 74; 124pp; English.

XX This invention relates to a novel method for preparing peptide linked
 CC oligomeric compounds by deprotecting the hydroxyl groups of a compound
 CC derivatising support medium, reacting deprotected hydroxyl groups with a
 CC nucleoside to form a compound from which a capped compound is formed,

oxidized and cleaved to form an oligomeric compound having a reactive sulfur moiety. The reactive sulphur moiety is reacted with peptide with functional group reactive with sulfur moiety, to form a peptide linked oligomeric compound. The method of the invention is useful for preparing an oligomeric compound. The oligomeric compounds can be used in diagnostics, therapeutics and as research reagents and kits. They can also be used in pharmaceutical compositions by including a suitable diluent or carrier. The oligomeric compounds of the invention can further be used for treating organisms having a disease characterised by the undesired production of a protein. This method is suitable for large scale synthesis of oligomeric compounds, the methods provide improved synthetic schemes which avoid the problem of prior art. The synthetic methods employed equimolar amounts of functionalised oligomeric compounds and peptide reagents which has successfully resulted in large scale synthesis. This scaled up synthesis is significantly larger than any synthesis method described previously. The methods are highly economical. The present sequence represents a peptide used in the creation of a peptide linked oligomeric compound of the invention

XX SQ Sequence 20 AA;

Query Match 83.1%; Score 98; DB 5; Length 20;
 Best Local Similarity 85.0%; Pred. No. 7.8e-07;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1 CXGGRQIKIWFQNRMKWK 20

RESULT 13
 ABU63051
 ID ABU63051 standard; protein; 20 AA.
 AC ABU63051;

DT 18-SEP-2003 (first entry)

DE Peptide linked oligomeric compound associated peptide #2.

KW Peptide linked oligomeric compound; diagnostic; therapeutic;
 KW research reagent; protein production inhibitor;
 KW phosphothioate 2'-O-MOE gapmer.

OS Synthetic.

FX Key Location/Qualifiers

FT Modified-site 2 /label= OTHER
 FT /note= "OTHER= Aminobutyric acid"

FT US2002156235-A1.

PN 24-OCT-2002.

PD 07-SEP-2001; 2001US-00949474.

PF 08-SEP-2000; 2000US-00659517.

PR (MANO/) MANOHARAN M.

PA (GUZA/) GUZAEV A P.

XX Manoharan M, Guzaev AP;

PI WPI; 2003-521518/49.

DR Preparation of peptide linked oligomeric compounds useful in diagnostics
 PT involves reacting deprotected hydroxyl group with nucleoside having a
 PT protected hydroxyl group and an activated phosphorus containing
 PT substituent group.

PS Example 20; Page 24; 45pp; English.

CC The invention describes preparation of a peptide linked oligomeric
 CC compound (I) involving preparing a deprotected hydroxyl group with a
 CC nucleoside having a protected hydroxyl group and activated phosphorus
 CC containing substituent group. (I) is useful in the preparation of peptide
 CC linked oligomeric compounds useful in diagnostics, therapeutics and as
 CC research reagents. They are useful in the treatment of diseases
 CC characterised by the undesired production of a protein in organisms such
 CC as bacteria, yeast, protozoa, algae, plants and animals including warm-
 CC blooded animals. The method is applicable to a large-scale synthesis of
 CC peptide-linked oligomeric compounds; is cost-effective and efficient. The
 CC method reduces the cost of preparation of (I). The process provides (I)
 CC without the problems of aggregation associated with electrostatic
 CC interactions. The process provides (I) with an increase in efficiency and
 CC provides an improved synthetic scheme avoiding the problems encountered
 CC during synthesis of cationic peptides e.g. problems generated by use of
 CC excess peptide reagents used. This is the amino acid sequence of a
 CC peptide created for use in the preparation of a peptide linker oligomeric
 CC compound

XX SQ Sequence 20 AA;

Query Match 83.1%; Score 98; DB 6; Length 20;
 Best Local Similarity 85.0%; Pred. No. 7.8e-07;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1 CXGGRQIKIWFQNRMKWK 20

RESULT 14
 AAY42291
 ID AAY42291 standard; protein; 60 AA.
 AC AAY42291;

DT 06-DEC-1999 (first entry)

DE Drosophila antennapodia homeobox domain peptide.

KW Cytochrome; targeting; localisation; cancer; tumour; prodrug; reduction;
 KW nucleus.

OS Drosophila sp.

OS Synthetic.

FX Key Location/Qualifiers

FT Misc-difference 4 /note= "Encoded by GGA"

FT WO9945127-A2.

PN 10-SEP-1999.

PD 05-MAR-1999; 99WO-GB000674.

PF 06-MAR-1998; 98GB-00004841.

PR 19-AUG-1998; 98GB-00018103.

PR 29-JAN-1999; 99GB-00002081.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Stratford IJ, Patterson AV, Kingsman SM, Kan O, Griffiths L;

PI Mitrophanous K;

XX WPI; 1999-551046/46.

DR N-PSDB; AAZ19783.

XX New prodrug activating agent targeted to selected cells or tissues,
 PT particularly hypoxic cells, for treating e.g. tumors.

PS Example 5; Fig 3; 187pp; English.

CC This sequence represents a Drosophila antennapedia homeobox domain
 CC peptide (pAntp), involved in transcriptional localisation. pAntp can be
 CC fused to cytochrome P450 reductase (P450R) derivatives such as anchorless
 CC P450R (AAV42287) or FN fragment (AAV42288). This enables the fusion
 CC protein to be delivered to other cells where it is then transported to
 CC the nucleus. Many drugs' sites of action are in the nucleus, rather than
 CC the cytoplasm, where P450R normally functions. P450R or its derivatives
 CC can be used to activate prodrugs to their active form via reduction.
 CC Administration of a prodrug is useful where the active drug may be
 CC metabolised before it reaches its site of action or where the active drug
 CC is cytotoxic, e.g., anticancer drugs. Targeted delivery of such prodrug
 CC activators allows a reduction in dose of the prodrug, and thus of
 CC systemic side-effects. P450R derivative fusion proteins, or vectors that
 CC express them, are specifically used to treat tumours, inflammation,
 CC atherosclerosis and muscular dystrophy, but may also be used to treat
 CC many other conditions, e.g., cerebral malaria, rheumatoid arthritis, or
 CC conditions associated with hypoxia, ischaemia or hypoglycaemia, or to
 CC deliver antibiotics, antiviral agents, analgesics, anaesthetics, anti-
 CC inflammatory, antineoplastic agents and diagnostic agents
 CC
 XX Sequence 60 AA;

Query Match 81.4%; Score 96; DB 2; Length 60;
 Best Local Similarity 85.0%; Pred. No. 4.2e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 CSSCRQIKIWFQNRMMKWK 20
 Db 39 CLTERQIKIWFQNRMMKWK 58

RESULT 15
 AAY04364
 ID AAY04364 standard; protein; 60 AA.
 AC AAY04364;

DT 21-JUN-1999 (first entry)

DE Drosophila antennapedia Antp protein homeodomain.

KW Drosophila; antennapedia; Antp gene; homeodomain; vaccine; cancer;
 KW geriatric disease; bacterial infection; viral infection.

OS Drosophila sp.

XX WO9911809-A1.

XX 11-MAR-1999.

XX 02-SEP-1998; 98WO-GB002628.

XX 02-SEP-1997; 97GB-00018609.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Crisanti A;

XX WPI; 1999-214708/18.

XX New conjugate comprising a homeodomain of Antennapedia - useful for
 XX prevention or treatment of cancer, geriatric diseases and bacterial or
 XX viral infections.

PS Example 1; Page 42; 45pp; English.

XX The present invention describes a conjugate comprising: (i) a first
 CC region comprising the homeodomain of antennapedia or a variant; and (ii)
 CC a second region not naturally associated with the first non-denatured
 CC region. The conjugate is useful in an expression system, and in
 CC pharmaceutical compositions e.g. vaccines for prevention or treatment of
 CC cancer, geriatric diseases, and bacterial or viral infections. It is also
 CC useful in gene therapy, protein sorting and for DNA synthesis. The

CC present sequence represents the Drosophila antennapedia (Antp) protein
 CC homeodomain
 XX
 SQ Sequence 60 AA;

Query Match 81.4%; Score 96; DB 2; Length 60;
 Best Local Similarity 85.0%; Pred. No. 4.2e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CSSCRQIKIWFQNRMMKWK 20
 Db 39 CLTERQIKIWFQNRMMKWK 58

RESULT 16
 AAY27403

ID AAY27403 standard; protein; 60 AA.

XX AAY27403;

XX 23-NOV-1999 (first entry)

XX D. melanogaster pAntp homeobox domain.

KW Prodrug; localization domain; tumor-selective antibody; cytochrome P450;
 KW prodrug activating domain; modified hematopoietic stem cell; MHSC; tumor;
 KW inflammation; atherosclerosis; muscular dystrophy; cerebral malaria;
 KW rheumatoid arthritis; hypoxia; ischemia; hypoglycemia; pAntp;
 KW antennapedia protein homeobox peptide.

OS Drosophila melanogaster.

XX Key Location/Qualifiers
 XX Misc-difference 4 /note= "encoded by GGA"

XX WO9945126-A2.

XX 10-SEP-1999.

XX 05-MAR-1999; 99WO-GB000672.

XX 06-MAR-1998; 98GB-00004841.

XX 19-AUG-1998; 98GB-00018103.

XX 29-JAN-1999; 99GB-00002081.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Stratford IJ, Patterson AV, Kingman SM, Kan O, Griffiths L;
 XX Mitrophanous K;

XX WPI; 1999-540852/45.

XX N-PSDB; AA207806.

XX New prodrug activating agent targeted to selected cells or tissues,
 XX particularly hypoxic cells, for treating e.g. tumors or inflammation.

PS Example 5; Fig 3C; 149pp; English.

XX The invention provides a new prodrug activating agent that comprises: (i)
 CC a localization domain (LD); other than a tumor-selective antibody) and a
 CC prodrug activating domain (PAD); (ii) at least one nucleic acid encoding
 CC a cytochrome P450 and under control of at least one constitutive or
 CC inducible expression control sequence or (iii) a modified hematopoietic
 CC stem cell (MHSC) containing at least one nucleic acid encoding a PAD and
 CC under control of elements as in (ii). The prodrug activating agent or
 CC vectors that express them, are specifically used to treat tumors,
 CC inflammation, atherosclerosis and muscular dystrophy, but may also be
 CC used to treat many other conditions, e.g. cerebral malaria, rheumatoid
 CC arthritis, or conditions associated with hypoxia, hypoglycemia or
 CC ischemia, or to deliver antibiotics, antiviral agents, analgesics,
 CC anesthetics, anti-inflammatory, antineoplastic agents and diagnostic
 CC agents. LD optimize activity of PAD, e.g. by delivering it to selected

CC locations or by delivering it to neighboring cells (bystander effect),
 CC and allow a reduction in dose of prodrug, and thus of systemic side-
 CC effects. Nucleic acids encoding the agent may be expressed selectively in
 CC hypoxic cells. The present sequence represents the D. melanogaster
 CC antennapedia protein homeobox peptide (pAntp). This is used in the
 CC construction of a fusion protein comprising pAntp and a human p450
 CC reductase derivative alP450R
 XX
 SQ Sequence 60 AA;
 Query Match 81.4%; Score 96; DB 2; Length 60;
 Best Local Similarity 85.0%; Pred. No. 4.2e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CSSCRQIKIWFQNRMKWK 20
 Db 39 CLTERQIKIWFQNRMKWK 58
 RESULT 17
 ABB84470
 ID ABB84470 standard; protein; 60 AA.
 XX
 AC ABB84470;
 XX
 DT 23-DEC-2002 (first entry)
 XX
 DE S. cerevisiae Antp homeobox region.
 XX
 KW Nuclear factor; transcription regulation; NF-kappaB; NF-kbeta; Ig gene;
 KW signalling; interleukin-1; tumour necrosis factor-alpha; antibacterial;
 KW nuclear translocation; virucide; infection; immunoglobulin.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US6410516-B1.
 XX
 PD 25-JUN-2002.
 XX
 PF 05-JUN-1995; 95US-00464364.
 XX
 PR 09-JAN-1986; 86US-00817441.
 PR 24-DEC-1986; 86US-00946365.
 PR 12-FEB-1988; 88US-00155207.
 PR 01-MAR-1988; 88US-00162680.
 PR 05-DEC-1988; 88US-00280173.
 PR 03-MAR-1989; 89US-00318901.
 PR 21-APR-1989; 89US-00341436.
 PR 13-NOV-1991; 91US-00791898.
 PR 06-APR-1995; 95US-00418266.
 XX
 (HARD) HARVARD COLLEGE.
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA (WHEE) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Baltimore D, Sen R, Sharp PA, Staudt L, Lebowitz JH;
 PI Baldwin AS, Clerc RG, Corcoran LM, Baeuerle PA, Lenardo MJ, Fan C;
 PI Maniatis TP;
 XX
 DR WPI; 2002-654437/70.
 XX
 XX Reduction of expression of genes transcriptionally regulated by nuclear
 XX transcriptional regulatory factor NF-kB useful e.g. to reduce effects of
 PT viral or bacterial infection on mammalian cells, by reducing NF-kB
 PT activity in the cell.
 PT
 XX Example 7; Fig 20; 106pp; English.
 PS
 XX This invention describes a novel method for inhibiting a gene whose
 CC transfection is regulated by NF-kappaB, a nuclear factor involved in the
 CC regulation of Ig (immunoglobulin) gene transcription (distribution).
 CC Inhibition can be achieved by reducing NF-kB signalling in the cell
 CC (optionally by reducing interleukin-1 or tumour necrosis factor-alpha

CC activity). A method is also described for reducing bacterial
 CC lipopolysaccharide-induced nuclear translocation of NF-kB in eukaryotic
 CC cells by inhibiting: (a) modification of new protein, tKB, which reduces
 CC binding to NF-kB; (b) degradation of tKB; or (c) dissociation of NF-kB-
 CC tKB complexes so as to reduce nuclear translocation of NF-kB in the
 CC cells. The products of the invention have virucide and antibacterial
 CC activity. The method is useful to inhibit expression of genes whose
 CC transcription is regulated by NF-kB, especially in mammalian (preferably
 CC human and especially immune or lymphoid) cells. It can be used to reduce
 CC expression of viral genes (e.g. cytomegalovirus, human immunodeficiency
 CC virus or simian virus 40 genes, especially to reduce the effects of viral
 CC infections on mammalian cells. It is also useful to reduce expression of
 CC cytokine genes or bacterial lipopolysaccharide-induced genes to reduce
 CC effects of bacterial lipopolysaccharides and/or bacterial infection on
 CC mammalian cells (especially immune cells). For example, bacterial
 CC lipopolysaccharide-induced expression of cytokines or tumor necrosis
 CC factor-alpha may be reduced so reducing the effects of bacterial
 CC lipopolysaccharides during infection. The method of inhibiting expression
 CC in a mammalian cell of a gene whose transcriptional activity is activated
 CC by binding of NF-kB to the gene, comprises introducing a nucleic acid
 CC decoy molecule into the cell in an amount which is sufficient to inhibit
 CC expression of the gene, the decoy contains a NF-kB binding site that
 CC binds to NF-kB. This sequence is used to illustrate the method described
 CC in the disclosure of the invention
 XX
 SQ Sequence 60 AA;

Query Match 81.4%; Score 96; DB 5; Length 60;
 Best Local Similarity 85.0%; Pred. No. 4.2e-06;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20
 Db 39 CLTERQIKIWFQNRMKWK 58

RESULT 18

AAE00811
 ID AAE00811 standard; peptide; 61 AA.
 XX

AC AAE00811;

XX 02-JUL-2001 (first entry)

DE D. melanogaster antennapedia (Antp) homeodomain transcription factor.

KW Fruit fly; Islet/Duodenum homeobox-1; IDX-1; antidiabetic; gene therapy;
 KW transactivation; somatostatin promoter; diabetes mellitus; homeodomain;
 KW transcription factor; antennapedia; Antp.

OS Drosophila melanogaster.

Key	Location/Qualifiers
FT Region	19..23 /label= Helix_1
FT Region	27..36 /label= Helix_2
FT Region	40..56 /label= Helix_3

PN US6210960-B1.

XX 03-APR-2001.

XX 19-NOV-1996; 96US-00751344.

XX 23-FEB-1994; 94US-00202044.

XX (GEHO) GEN HOSPITAL CORP.

XX Habener JF, Miller CP;

XX WPI; 2001-280863/29.

XX New IDX-1 polypeptides and polynucleotides encoding the IDX-1
PT polypeptides, useful for treating or preventing diabetes mellitus type I
PT or II, or symptoms associated with diabetes.
XX
PS Example 1; Fig 1C; 42pp; English.
XX
CC The patent discloses IDX-1 sequence, a novel homeodomain transcription
CC factor which is present in pancreatic duct cells, but not in pancreatic
CC alpha-cells. IDX-1 is expressed in pancreatic islets and ducts, and in
CC the duodenum, and binds to and activates transcription and regulatory
CC sequence elements in the 5' flanking region of the rat somatostatin gene.
CC IDX-1 regulates the transcription of somatostatin gene. The IDX-1
CC polypeptides and polynucleotides are useful for treating or preventing
CC diabetes mellitus type I or II, and symptoms or conditions associated
CC with diabetes. IDX-1 is used to immunise animals to identify antibodies
CC that bind to other domains of IDX-1 or to screen for antibodies that
CC specifically bind to the promoter binding domain of IDX-1. It is also
CC used in gene therapy. The diagnostic kit is useful for detecting a defect
CC in a gene encoding IDX-1. The present sequence is a homeodomain
CC transcription factor of antennapedia (Antp) gene from Drosophila
CC melanogaster. Antp sequence was used for the isolation and sequence
CC analysis of the rat IDX-1 cDNA
XX
SQ Sequence 61 AA;

Query Match 81.4%; Score 96; DB 4; Length 61;
Best Local Similarity 85.0%; Pred. No. 4.3e-06;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMKWK 20
| : |||||
Db 40 CLTERQIKIWFQNRMKWK 59

RESULT 19
AAE32060
ID AAE32060 standard; protein; 64 AA.
XX
AC AAE32060;
XX
DT 24-MAR-2003 (first entry)
XX
DE Antennapedia C3APL peptide.
XX
XX Drug delivery construct; axon growth; nerve injury; ischaemic damage;
KW stroke injury; gene therapy; neuroleptic; neuroprotective.
XX
XX Antennapedia.
OS
XX WO200283179-A2.
XX
XX 24-OCT-2002.
XX
XX 08-APR-2002; 2002WO-CA000480.
XX
XX 12-APR-2001; 2001CA-02342970.
PR
XX 13-NOV-2001; 2001CA-02362004.
PR
XX 15-JAN-2002; 2002CA-02367636.
PR
XX (BIOA-) BIOAXONE THERAPEUTIQUE INC.
PA
XX
XX McKerracher L;
PI
XX WPI; 2003-092963/08.
DR
XX New drug delivery construct comprising a transport and active agent
PT region, useful for the manufacture of a pharmaceutical composition for
PT treating nerve injury.
PT
XX Claim 46; Page 64; 189pp; English.
PS
XX The invention relates to a new drug delivery construct comprises at least

CC one transport agent region and an active agent region. The transport
CC agent region is able to facilitate the uptake of the active agent region
CC into a cell. The active agent region is an active therapeutic agent
CC region able to facilitate axon growth and an analogue. The drug delivery
CC construct is useful for suppressing the inhibition of neuronal axon
CC growth, facilitating axon growth, treating nerve injury, treating
CC ischaemic damage related to stroke injury. The drug delivery construct
CC and the drug conjugate are useful for the manufacture of a pharmaceutical
CC composition for treating nerve injury. The invention is useful in gene
CC therapy. The present sequence is Antennapedia C3APL peptide
XX
SQ Sequence 64 AA;

Query Match 81.4%; Score 96; DB 6; Length 64;
Best Local Similarity 85.0%; Pred. No. 4.5e-06;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMKWK 20
| : |||||
Db 43 CLTERQIKIWFQNRMKWK 62

RESULT 20
ABB99729
ID ABB99729 standard; protein; 128 AA.
XX
AC ABB99729;
XX
DT 24-MAR-2003 (first entry)
XX
DE Amino acid sequence of AntpHD-Cw3 protein.
XX
XX Lipid-based vehicle; internalizing peptide; immune response; cancer;
KW cardiovascular disease; hypertension; cardiac arrhythmia; restenosis;
KW infection; inflammation; autoimmune disease; vaccine; AntpHD-Cw3;
KW Antp homeodomain; Cw3.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
PH Peptide 3..12 "His tag"
FT Peptide /note= "His tag"
FT Peptide 24..84
FT Peptide /note= "AntpHD"
FT Peptide 87..96
FT Peptide /note= "C-myc tag"
FT Peptide 104..113
FT Peptide /note= "HLA-Cw3"
FT Peptide 119..128
FT Peptide /note= "C-myc tag"
XX
XX WO200298465-A2.
XX
XX 12-DEC-2002.
XX
XX 07-JUN-2002; 2002WO-CA000853.
XX
XX 07-JUN-2001; 2001US-0296158P.
PR
XX (CELA-) CELATOR TECHNOLOGIES INC.
PA
XX
XX Bally MB, Schutze-Redelmeier M, Chikh G;
PI
XX WPI; 2003-140559/13.
DR
XX
XX Composition for delivery of a biologically active agent to a cell for
PT treating cancer, cardiovascular or autoimmune diseases, comprises an
PT internalizing peptide/agent complex conjugated to a lipid-based vehicle.
PT
XX Example; Fig 1B; 33pp; English.
PS
XX The specification describes a composition for delivery of a biologically
CC active agent to a cell. The composition comprises an internalizing

CC peptide/agent complex conjugated to a lipid-based vehicle. By conjugating
CC the internalizing peptide/agent complex to a lipid-based vehicle, greater
CC loading of the complex to the vehicle is achieved, and greater stability
CC is achieved during delivery to a target cell than is the case resulting
CC from hydrophobic interaction of the complex with a lipid-based vehicle.
CC Despite the presence of a strong bond between the complex and the
CC vehicle, the complex is still able to enter the cytosol of the target
CC cell at highly efficient rates. The composition is useful for delivery of a
CC an agent to a cell or for preparation of a medicament for treatment of a
CC patient. The composition is useful for bringing about a desired
CC biological effect such as immune response or for delivery of agents for
CC treatment of a variety of diseases such as cancer, cardiovascular
CC diseases (such as hypertension, cardiac arrhythmia, restenosis),
CC bacterial, viral, fungal or parasitic infections, inflammation or
CC autoimmune diseases, or as a vaccine. The present sequence represents a
CC protein, designated AntpHD-Cw3. This peptide comprises the Antp
CC homeodomain (AntpHD) as the internalising peptide, as well as residues
CC from Cw3. The peptide was conjugated to liposomes to produce a
CC composition of the invention, and used to induce an immune response in
CC mice against Cw3
XX
XX
SQ Sequence 128 AA;
Query Match 81.4%; Score 96; DB 6; Length 128;
Best Local Similarity 85.0%; Pred. No. 8.6e-06;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSSCROIKWIFONRRMKWK 20
Db 63 CLTERQIKWIFONRRMKWK 82
RESULT 21
AAE10922
ID AAE10922 standard; protein; 217 AA.
XX
AC AAE10922;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human HOXB7 variant protein.
XX
KW Human; cytostatic; cancer-related antigen; homeobox protein; HOXA7;
KW HOXB7; ADP ribosylation factor 1; Arf-1; ATP dependent iron transporter;
KW ABC-7; neoplastic process; ovarian cancer; benign serous cystadenoma;
KW vaccine; variant.
XX
OS Homo sapiens.
XX
PN WO20016853-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-US007896.
XX
XX 14-MAR-2000; 2000US-0189226P.
PR 28-DEC-2000; 2000US-0258452P.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Roden R, Naora H;
XX
DR WPI; 2001-596909/67.
DR N-PSDB; AAD18279.
XX
PT Novel cancer-related antigen useful for prognosis, diagnosis and
PT treatment of cancer, especially ovarian cancer in an individual,
PT comprises a fragment isolated from bacteriophage lambda.
XX
PS Claim 19; Fig 12; 67pp; English.
XX
CC The patent discloses autoantibodies in cancer patients specific for novel
CC cancer-related antigens that are normally intracellular including

CC homeobox proteins, HOXA7, HOXB7, ADP ribosylation factor 1 (Arf-1), ATP
CC dependent iron transporter ABC-7 and a novel protein encoded by
CC EcoRI/XhoI fragment isolated from bacteriophage lambda clone 44B.1. The
CC presence of these autoantibodies is correlated with neoplastic processes
CC in patients. Proteins of the invention are useful for screening for
CC cancer in an individual. HOXB7 is useful for screening for cancer other
CC than breast cancer, renal cell carcinoma, colon cancer or melanoma in an
CC individual, by determining whether cells in the individual are expressing
CC a gene product of HOXB7, expression of which is correlated with increased
CC likelihood of cancer in the individual. It is useful for screening
CC ovarian cancer or benign serous cystadenoma. HOXB7 proteins are useful to
CC distinguish between neoplastic and non-neoplastic fluid accumulations in
CC patients carrying a malignant diagnosis and in screening methods for
CC therapeutically active materials. HOXB7 antibodies are useful for
CC detecting epitopes found on proteins of the invention in histological
CC sections of ovarian cancer tissue as well as in other solid tumours such
CC as breast cancer and melanoma. The proteins of the invention are also
CC used as vaccines. The present sequence is HOXB7 variant protein from
CC human
XX
XX SQ Sequence 217 AA;
Query Match 81.4%; Score 96; DB 4; Length 217;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSSCROIKWIFONRRMKWK 20
Db 175 CLTERQIKWIFONRRMKWK 194
RESULT 22
ABG96337
ID ABG96337 standard; protein; 217 AA.
XX
AC ABG96337;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M448.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nontuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Wieby PO, Mills GB;
PI Bast RC, Lu X, Schmandt RE, Zhao X, Glatt K;
XX
XX WPI; 2002-723277/78.
DR N-PSDB; ABS76433.
DR

XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.

XX Disclosure, Page 248-249; 481pp; English.

XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterizing cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention

XX Sequence 217 AA;
SQ Query Match 81.4%; Score 96; DB 5; Length 217;
Best Local Similarity 85.0%; Pred. No. 1.4e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
Db 175 CLTERQIKIWFQNRMRKWK 194

RESULT 23
AA58610
ID AAY58610 standard; protein; 230 AA.
XX AC AAY58610;
XX 11-APR-2000 (first entry)
XX Protein regulating gene expression PRGE-3.
XX Protein regulating gene expression; PRGE-3; human; cancer; trauma;
KW anticancer; antitumour; antitraumatic; therapy; diagnosis.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 35
FT /note= "O-phosphorylated"
FT Peptide 119..124
FT /note= "signature sequence"
FT Modified-site 126
FT /note= "O-phosphorylated"
FT Peptide 131..188
FT /note= "signature sequence"
FT Peptide 163..186
FT /note= "signature sequence"
FT Modified-site 170

FT /note= "O-phosphorylated"
XX WO9964596-A2.
PN 16-DEC-1999.
XX 11-JUN-1999; 99WO-US013281.
XX 12-JUN-1998; 98US-0089029P.
PR 29-JUL-1998; 98US-0094575P.
PR 14-OCT-1998; 98US-0104624P.
XX (INCY-) INCYTE PHARM INC.
XX Lal P, Yue H, Tang YT, Hillman JL, Bandman O, Corley NC;
PI Guegler KJ, Gorgone GA, Baughn MR, Patterson C, Lu DAM;
XX WPI; 2000-116543/10.
DR N-PSDB; AAZ57841.
XX New human polypeptides that regulate gene expression, for treatment,
PT prevention and diagnosis of, e.g. cancer.
XX Claim 1; Page 84; 150pp; English.
XX The present sequence is that of new human protein regulating gene
CC expression PRGE-3. It was deduced from Incyte clone 996352 obtained from
CC a kidney tumour CDNA library. PRGE-3 is characterised as a homeodomain
CC protein. It is expressed in reproductive, cardiovascular and urologic
CC tissues associated with cancer and trauma diseases, disorders or
CC conditions. The invention provides PRGE polypeptides (see AAY58608-38)
CC and polynucleotides (see AAZ57839-69), expression vectors, host cells,
CC antibodies, agonists and antagonists. It also provides methods for
CC diagnosing, treating or preventing disorders associated with expression
CC of PRGE

XX Sequence 230 AA;
SQ Query Match 81.4%; Score 96; DB 3; Length 230;
Best Local Similarity 85.0%; Pred. No. 1.5e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
Db 168 CLTERQIKIWFQNRMRKWK 187

RESULT 24
AAE32047
ID AAE32047 standard; protein; 295 AA.
XX AC AAE32047;
XX 24-MAR-2003 (first entry)
XX Clostridium botulinum and Antennapedia C3APL protein.
XX Drug delivery construct; axon growth; nerve injury; ischaemic damage;
KW stroke injury; gene therapy; neuroleptic; neuroprotective;
KW fusion protein.
XX Clostridium botulinum.
OS Antennapedia.
OS Chimeric.
XX Key Location/Qualifiers
FH Region 1..231
FT /note= "ADP-ribosyl transferase C3"
FT Region 232..295
FT /note= "Antennapedia protein"
XX WO200283179-A2.
PN
XX

PD 24-OCT-2002.
XX
XX
PF 08-APR-2002; 2002WO-CA000480.
XX
XX PR 12-APR-2001; 2001CA-02342970.
XX PR 13-NOV-2001; 2001CA-02362004.
XX PR 15-JAN-2002; 2002CA-02367636.
XX
XX (BIOA-) BIOAXONE THERAPEUTIQUE INC.
XX
XX PI McKerracher L;
XX
XX DR WPI; 2003-092963/08.
XX DR N-PSDB; AAD49471.
XX
XX PT New drug delivery construct comprising a transport and active agent
XX PT region, useful for the manufacture of a pharmaceutical composition for
XX PT treating nerve injury.
XX
XX PS Claim 41; Page 96-97; 188pp; English.
XX
XX CC The invention relates to a new drug delivery construct comprises at least
XX CC one transport agent region and an active agent region. The transport
XX CC agent region is able to facilitate the uptake of the active agent region
XX CC into a cell. The active agent region is an active therapeutic agent
XX CC region able to facilitate axon growth and an analogue. The drug delivery
XX CC construct is useful for suppressing the inhibition of neuronal axon
XX CC growth, facilitating axon growth, treating nerve injury, treating
XX CC ischemic damage related to stroke injury. The drug delivery construct
XX CC and the drug conjugate are useful for the manufacture of a pharmaceutical
XX CC composition for treating nerve injury. The invention is useful in gene
XX CC therapy. The present sequence is C3APL protein comprising ADP-ribosyl
XX CC transferase C3 protein and antenapedia protein
XX
XX SQ Sequence 295 AA;

Query Match 81.4%; Score 96; DB 6; Length 295;
Best Local Similarity 85.0%; Pred. No. 1.9e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMRKWK 20
| : |||||
Db 274 CLTERQIKIWFQNRMRKWK 293

RESULT 25
ABB72035
ID ABB72035 standard; protein; 378 AA.
XX
XX AC ABB72035;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 42897.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX FN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX

DR WPI; 2001-656860/75.
DR N-PSDB; ABL16138.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Disclosure; SEQ ID NO 42897; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16178-ABL30511); expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 378 AA;

Query Match 81.4%; Score 96; DB 4; Length 378;
Best Local Similarity 85.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMRKWK 20
| : |||||
Db 335 CLTERQIKIWFQNRMRKWK 354

Search completed: May 24, 2004, 17:22:22
Job time : 55.5135 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:18:06 ; Search time 11.6216 Seconds
(without alignments)
165.539 Million cell updates/sec

Title: US-09-977-349-3

Perfect score: 118

Sequence: 1 CSSCRQIKIWFQNRMRKWKX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 78:**

1: Piri:**

2: Pir2:**

3: Pir3:**

4: Pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	99	83.9	45	2	PC1216		homeotic protein D
2	99	83.9	76	2	C43559		homeo box protein
3	99	83.9	78	2	I51342		homeotic protein H
4	99	83.9	82	2	S08302		homeotic protein Z
5	99	83.9	86	2	JT0489		homeotic protein T
6	99	83.9	107	2	B61045		homeotic protein H
7	99	83.9	118	2	A24777		homeotic protein H
8	99	83.9	118	2	JT0273		homeotic protein M
9	99	83.9	118	2	B24777		homeotic protein b
10	99	83.9	138	2	S20087		homeobox A5 protein
11	99	83.9	148	2	PC4071		homeotic protein H
12	99	83.9	158	2	A27348		homeotic protein H
13	99	83.9	230	2	S00592		homeotic protein H
14	99	83.9	233	2	I65197		homeotic protein H
15	99	83.9	236	1	S09256		homeotic protein H
16	99	83.9	245	1	S10092		hypothetical prote
17	99	83.9	246	2	T46446		homeotic protein H
18	99	83.9	250	1	A31757		homeotic protein H
19	99	83.9	251	1	B60492		homeotic protein H
20	99	83.9	255	1	WJHU4B		homeotic protein H
21	99	83.9	264	1	WJHU3E		homeotic protein H
22	99	83.9	264	1	S35219		homeotic protein H
23	99	83.9	269	1	A43551		homeotic protein H
24	99	83.9	269	2	A45578		homeotic protein H
25	99	83.9	270	1	WJHULC		homeotic protein H
26	99	83.9	270	1	WJMS13		homeotic protein H
27	99	83.9	275	1	WJZFX2		homeotic protein R
28	98	83.1	67	2	B27471		homeotic protein H
29	98	83.1	85	2	I65198		homeotic protein H

30	83.1	85	2	A25472	homeotic protein H
31	83.1	155	2	S25846	homeotic protein H
32	83.1	285	1	A43556	homeotic protein H
33	83.1	309	1	S09257	homeotic protein H
34	83.1	320	1	A39724	homeotic protein H
35	81.4	33	2	S57235	antennapedia prote
36	81.4	42	2	I65241	homeotic protein H
37	81.4	48	2	I51439	homeobox protein -
38	81.4	66	2	S15536	homeotic protein H
39	81.4	66	2	S15538	homeotic protein H
40	81.4	71	2	JC1161	homeotic protein 3
41	81.4	71	2	A60084	homeotic protein H
42	81.4	74	2	D34510	homeo box protein
43	81.4	75	2	I51341	homeotic protein S
44	81.4	75	2	S58852	homeotic protein H
45	81.4	81	2	S47605	homeotic protein H
46	81.4	83	2	S47603	homeotic protein H
47	81.4	83	2	S50066	homeotic protein H
48	81.4	86	2	A34510	homeotic protein H
49	81.4	86	2	S08303	homeotic protein H
50	81.4	87	2	S00589	homeotic protein M
51	81.4	88	2	A03317	homeotic protein z
52	81.4	96	2	S08639	homeotic protein H
53	81.4	96	2	A05266	homeotic protein H
54	81.4	97	2	C27176	homeotic protein m
55	81.4	97	2	A24779	homeotic protein H
56	81.4	103	2	A32167	homeotic protein R
57	81.4	105	2	A27471	homeotic protein s
58	81.4	106	2	S36448	homeobox protein -
59	81.4	113	2	T10775	homeotic protein m
60	81.4	119	2	A03314	homeotic protein H
61	81.4	153	1	WJHU3C	homeotic protein H
62	81.4	153	1	WJMSX6	homeotic protein H
63	81.4	209	2	A43553	homeotic protein H
64	81.4	217	1	WJHU2C	homeotic protein H
65	81.4	217	1	WJMSX2	homeotic protein H
66	81.4	220	2	A37371	Hox 2 protein type
67	81.4	220	2	S01063	homeotic protein H
68	81.4	222	2	S20029	homeotic protein H
69	81.4	224	2	S26400	homeotic protein H
70	81.4	224	2	A31324	homeotic protein H
71	81.4	228	2	S32563	homeotic protein H
72	81.4	229	2	A28329	homeotic protein H
73	81.4	232	2	S48125	homeotic protein H
74	81.4	234	1	S02014	homeotic protein H
75	81.4	234	2	S00992	homeotic protein H
76	81.4	235	2	A56568	homeotic protein H
77	81.4	235	2	S72429	homeotic protein H
78	81.4	242	1	A39164	homeotic protein a
79	81.4	295	2	S58850	homeotic protein A
80	81.4	378	2	A25399	homeotic protein A
81	81.4	394	2	S26492	homeotic protein S
82	81.4	413	1	S03631	homeotic protein H
83	79.7	81	2	B29585	homeotic protein H
84	79.7	250	1	A36170	antennapedia-like
85	78.0	60	2	PC2399	homeotic protein H
86	78.0	60	2	PC2400	homeotic protein H
87	78.0	105	2	S47602	homeotic protein H
88	78.0	227	2	S50067	homeotic protein D
89	78.0	232	1	A25108	homeotic protein (
90	78.0	590	2	A26638	homeotic protein H
91	77.1	50	2	S10907	homeotic protein H
92	77.1	60	2	I51437	homeotic protein H
93	77.1	62	2	B32391	homeotic protein H
94	77.1	66	2	S15534	homeotic protein m
95	77.1	69	2	B3785	homeotic protein H
96	77.1	70	2	B37042	homeotic protein H
97	77.1	81	2	A25180	homeotic protein H
98	77.1	95	2	B32830	homeotic protein H
99	77.1	96	2	A55278	homeotic protein H
100	77.1	99	2	A41605	homeotic protein H

ALIGNMENTS

```

RESULT 1
PC1216
homeotic protein DthbX1 - planarian (Dugesia tigrina) (fragment)
C:Species: Dugesia tigrina
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Oct-1997
C/Accession: PC1216
R/Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlich
Gene 121, 337-342, 1992
A:Title: Homeoboxes in flatworms.
A:Reference number: JCI1386; MUID:93077050; PMID:1359988
A:Accession: PC1216
A:Molecule type: DNA
A:Residues: 1-45 <ORI>
A:Cross-references: EMBL:X66822
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:1-45/Domain: homeobox homology (fragment) <HOX>

Query Match      83.9%; Score 99; DB 2; Length 45;
Best Local Similarity 90.0%; Pred. No. 3.2e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C S S C R Q I K I W F Q N R R M K W K 20
    | | | | | | | | | | | | | | | | | |
Db 26 C L S E R Q I K I W F Q N R R M K W K 45

RESULT 2
C43559
homeotic protein R3 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 24-Sep-1999
C/Accession: C43559
R/Falzon, M.; Chung, S.Y.
Development 103, 601-610, 1988
A:Title: The expression of rat homeobox-containing genes is developmentally regulated an
A:Reference number: A43559; MUID:89231502; PMID:2907739
A:Accession: C43559
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <PAL>
A:Cross-references: GB:M37567; NID:g204634; PIDN:AAA41343.1; PID:g204635
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:20-76/Domain: homeobox homology (fragment) <HOX>

Query Match      83.9%; Score 99; DB 2; Length 76;
Best Local Similarity 90.0%; Pred. No. 4.8e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C S S C R Q I K I W F Q N R R M K W K 20
    | | | | | | | | | | | | | | | | | |
Db 49 C L S E R Q I K I W F Q N R R M K W K 68

RESULT 3
I51342
homeo box protein - Atlantic salmon (fragment)
C:Species: Salmo salar (Atlantic salmon)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C/Accession: I51342
R/Fjose, A.; Molven, A.; Eiken, H.G.
Gene 62, 141-152, 1988
A:Title: Molecular cloning and characterization of homeobox-containing genes from Atlan
A:Reference number: I51341; MUID:88226009; PMID:2897318
A:Accession: I51342
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-78 <FO>
A:Cross-references: GB:M18904; NID:g213799; PIDN:AAA49560.1; PID:g213800

```

C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:2-58/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 78;
Best Local Similarity 90.0%; Pred. No. 4.9e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C S S C R Q I K I W F Q N R R M K W K 20

Db 39 C L S E R Q I K I W F Q N R R M K W K 58

RESULT 4

S08302
homeotic protein Hox 2.1 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 02-Dec-1993 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C/Accession: S08302
R/Wedden, S.B.; Pang, K.; Eichele, G.
Development 105, 639-650, 1989
A:Title: Expression pattern of homeobox-containing genes during chick embryogenesis.
A:Reference number: S08302; MUID:90126373; PMID:2575515
A:Accession: S08302
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-82 <MED>
A:Cross-references: EMBL:X16846; NID:g62905; PIDN:CAA34743.1; PID:g1334633
A:Note: the authors translated the codon GGA for residue 4 as Arg
C:Superfamily: homeotic protein Hox A5; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:8-64/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 82;
Best Local Similarity 90.0%; Pred. No. 5.1e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C S S C R Q I K I W F Q N R R M K W K 20

Db 45 C L S E R Q I K I W F Q N R R M K W K 64

RESULT 5

JT0489
homeotic protein ZF-13 - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 17-Oct-1997
C/Accession: JT0489
R/Njoelstad, P.R.; Molven, A.; Eiken, H.G.; Fjose, A.
Gene 73, 33-46, 1988
A:Title: Structure and neural expression of a zebrafish homeobox sequence.
A:Reference number: JT0489; MUID:89211958; PMID:2468579
A:Accession: JT0489
A:Molecule type: DNA
A:Residues: 1-86 <NUO>
C:Genetics:
A:Gene: ZF-13
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:11-67/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 86;
Best Local Similarity 90.0%; Pred. No. 5.3e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 C S S C R Q I K I W F Q N R R M K W K 20

Db 48 C L S E R Q I K I W F Q N R R M K W K 67

RESULT 6

B61045
homeotic protein TghBox 6 - sea urchin (Tripneustes gratilla) (fragment)

C;Species: Triboneustes gratillia
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
 C;Accession: B61045
 R;Wang, G.V.L.; Dolecki, G.J.; Carlos, R.; Humphreys, T.
 Dev. Genet. 11, 77-87, 1990
 A;Title: Characterization and expression of two sea urchin homeobox gene sequences.
 A;Reference number: A61045; MUID:90298585; PMID:1972915
 A;Accession: B61045
 A;Status: Preliminary
 A;Molecule type: DNA

A;Residues: 1-107 <WAN>
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;9-65/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 107;
 Best Local Similarity 90.0%; Pred. No. 6.3e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
 DB 46 CLSERQIKIWFQNRMRKWK 65

RESULT 7

A24777
 homeotic protein Hox B5 - human (fragment)
 N;Alternate names: homeotic protein Hox 2A; homeotic protein Hxl
 C;Species: Homo sapiens (man)
 C;Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 24-Sep-1999
 C;Accession: A24777; A25181; S15540; A05265
 R;Hauser, C.A.; Joyner, A.L.; Klein, R.D.; Learned, T.K.; Martin, G.R.; Tjian, R.
 Cell 43, 19-28, 1985
 A;Title: Expression of homologous homeo-box-containing genes in differentiated human ter
 A;Reference number: A24777; MUID:86079489; PMID:4075393
 A;Accession: A24777
 A;Molecule type: DNA
 A;Residues: 1-118 <HAU>
 A;Cross-references: EMBL:K02572; NID:G184279; PIDN:AAA52681.1; PID:G386781
 R;Simone, A.; Mavilio, F.; Bottero, L.; Giampaolo, A.; Russo, G.; Faiella, A.; Boncinelli
 Nature 320, 763-765, 1986
 A;Title: A human homeo box gene specifically expressed in spinal cord during embryonic
 A;Reference number: A25181; MUID:86203637; PMID:3453105
 A;Accession: A25181
 A;Molecule type: mRNA

A;Residues: 43-118 <SIM>
 A;Cross-references: EMBL:X03794; NID:G32363; PIDN:CAA27420.1; PID:G32364
 R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Stc
 Genome 31, 745-756, 1989
 A;Title: Organization of human class I homeobox genes.
 A;Reference number: S15036; MUID:90215256; PMID:2576652
 A;Accession: S15540
 A;Molecule type: DNA
 A;Residues: 43-108 <BON>
 C;Genetics:
 A;Gene: GDB:H0XB5
 A;Cross-references: GDB:120658; OMIM:142960
 A;Map position: 17q21.3-17q21.3
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;44-100/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 118;
 Best Local Similarity 90.0%; Pred. No. 6.8e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
 DB 81 CLSERQIKIWFQNRMRKWK 100

RESULT 8

JT0273

homeotic protein Hox 2A homolog - sheep (fragment)
 N;Alternate names: homeotic protein Obox-2-1
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 17-Oct-1997
 C;Accession: JT0273
 R;Choi, C.L.; Hudson, P.; Srauder, A.; Pietersz, G.; Brandon, M.
 Gene 63, 187-197, 1988
 A;Title: Molecular cloning and characterization of ovine homeo-box-containing genes.
 A;Reference number: JT0273; MUID:88255866; PMID:2898416
 A;Accession: JT0273
 A;Molecule type: DNA
 A;Residues: 1-118 <CHO>
 A;Note: the authors translated the codon CAG for residue 67 as Glu and TTG for residue 6.
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;44-100/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 118;
 Best Local Similarity 90.0%; Pred. No. 6.8e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
 DB 81 CLSERQIKIWFQNRMRKWK 100

RESULT 9

B24777
 homeotic protein Mui - mouse (fragment)
 C;Species: Mus musculus (house mouse)
 C;Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 17-Oct-1997
 C;Accession: B24777
 R;Hauser, C.A.; Joyner, A.L.; Klein, R.D.; Learned, T.K.; Martin, G.R.; Tjian, R.
 Cell 43, 19-28, 1985
 A;Title: Expression of homologous homeo-box-containing genes in differentiated human ter
 A;Reference number: A24777; MUID:86079489; PMID:4075393
 A;Accession: B24777
 A;Molecule type: DNA
 A;Residues: 1-118 <HAU>
 C;Genetics:
 A;Map position: 11
 C;Superfamily: unassigned homeobox proteins; homeobox homology
 C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F;44-100/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 118;
 Best Local Similarity 90.0%; Pred. No. 6.8e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
 DB 81 CLSERQIKIWFQNRMRKWK 100

RESULT 10

S20087
 homeotic protein box6 - sea urchin (Parechinus angulosus) (fragment)
 C;Species: Parechinus angulosus (angulate urchin)
 C;Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 05-Dec-1998
 C;Accession: S20087; PS0274
 R;Pfeffer, P.L.; von Holt, C.
 Gene 108, 219-226, 1991
 A;Title: Stage- and adult tissue-specific expression of a homeobox gene in embryo and ad

A;Reference number: PS0274; MUID:92084139; PMID:1684167
 A;Accession: S20087
 A;Molecule type: DNA
 A;Residues: 1-138 <PFE>
 A;Cross-references: EMBL:X54494
 A;Accession: PS0274
 A;Molecule type: DNA
 A;Residues: 51-138 <PFF>
 A;Cross-references: EMBL:X54494
 C;Genetics:

A;Gene: box6
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;60-115/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 138;
Best Local Similarity 90.0%; Pred. No. 7.7e-07;
Matches 18; Conservative 0; Mismatches 2; Indels

Qy 1 CSSCRQIKIWFQNRMMWK 20
pb 96 CLSEROIKIWFQNRMMWK 115

RESULT 11

PC4071
homeobox A5 protein - axolotl (fragment)
C/Species: Ambystoma mexicanum (axolotl)
C/Date: 15-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 23-Sep-2002
C/Accession: PC4071
R/Gaur, A.; Lemanski, L.F.; Dube, D.K.
Gene 122, 249-253, 1995
A/Title: Identification and expression of a homologue of the murine HoxA5 9
A/Reference number: PC4071; MUID:96032352; PMID:7557438
A/Accession: PC4071
A/Molecule type: mRNA
A/Residues: 1-148 <GAU>
A/Cross-references: GB:U19238; NID:G746417; PIDN:AAA91634.1; PID:G746418
C/Genetics:
A/Gene: HoxA5
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:74-130/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 2; Length 148;
Best Local Similarity 90.0%; Pred. No. 8.2e-07;
Matches 18; Conservative 0; Mismatches 2; Indels

QY 1 CSSCRQIKIWFO NRRMKWK 20
| | | | | | | | | |
pb 111 CLSEROIKIWFO NRRMKWK 130

Page 12

[illegible]

Query Match	83.9%;	Score 99;	DB 2;	Length 158;
Best Local Similarity	90.0%;	Pred. No. 8.6e-07;		
Matches	18: Conservative	0: Mismatches	2: Indels	

Qy 1 CSSCRQIKIWFO~~NR~~RMKWCK 20
pb 73 CLSEROIKIWFO~~NR~~RMKWCK 92

RESULT 13

S00592
homeotic protein Hox 4 - African clawed frog (fragment)
C/Species: Xenopus laevis (African clawed frog)
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 20-Jun-2000
C/Accession: S00592
A/Title: Xenopus homeobox-containing cDNAs expressed in early development.
A/Reference number: S00589; MUID:86157707; PMID:2894634
A/Accession: S00592
A/Molecule type: mRNA
A/Residues: 1-230 <R1>
A/Cross-references: EMBL:X07104; NID:G64755; PIN:CAA30125.1; PTD:gl334652
C/Superfamily: homeotic protein Hox A5; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
P:156-212/Domain: homeobox homology <Hox>

Query Match 83.9%; Score 99; DB 2; Length 230;
Best Local Similarity 90.0%; Pred. No. 1.2e-06;
Matches 18: Conservative 0; Mismatches 2; Indels

Qy 1 CSSCRQIKIWFQNRMMKWK 20
| | | | | | | | | |
Db 193 CLSERQIKIWFQNRMMKWK 212

RESUIT 14

765197
hox1.3 protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 22-Mar-2002
C;Accession: I65197
R;Gorski, D.H.; LePage, D.F.; Walsh, K.
BioTechniques 16, 856-858, 1994
A;Title: Cloning and sequence analysis of homeobox transcription factor cDN
A;Reference number: I52196; MUID:94347374; PMID:7915120
A;Accession: I65197
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-233 <RES>
A;Cross-references: GB:I03556; NID:g204643; PIDN:AAA67844.1; PID:g204644
C;Genetics:

A:Gene: *hoxl.3*
C:Superfamily: homeotic protein Hox A5; homeobox homology
C:Superfamily: DNA binding; homeobox; nucleus; transcription regulation
F:175-231/Domain: homeobox homology <Hox>

Query Match 83.9%; Score 99; DB 2; Length 233;
Best Local Similarity 90.0%; Pred. No. 1.2e-06;
Matches 18: Conservative 0; Mismatches 2; Indels

QY 1 CSSCQIKIWFO NRRMKWK 20
| | | | | | | | | |
db 212 CLSEROIKIWFO NRRMKWK 231

RESULT 15

S09256
 homeotic protein Hox D4 - chicken
 N:Alternate names: homeotic protein Chox-4.2; homeotic protein Chox-a
 C:Species: Gallus gallus (chicken)
 C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 24-Jul-1997
 C:Accession: S09256
 R:Sasaki, H.; Yokoyama, E.; Kuroiwa, A.
 Nucleic Acids Res. 18, 1739-1747, 1990
 A:Title: Specific DNA binding of the two chicken deformed family homeodomain

A;Reference number: S09256; MUID:90245562; PMID:1970866

A;Accession: S09256

A;Molecule type: mRNA

A;Residues: 1-236 <SAS>

A;Cross-references: EMBL:X52671; EMBL:X52672

C;Genetics:

A;Gene: hoxd-4

C;Function:

A;Description: control of embryonic development by tissue- and stage-specific regulation

C;Superfamily: homeotic protein Hox D4; homeobox homology

C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F;146-202/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 1; Length 236;

Best Local Similarity 90.0%; Pred. No. 1.2e-06;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20

DB 183 CLSERQIKIWFQNRMKWK 202

RESULT 16

S10092

homeotic protein Hox B4 - chicken

N;Alternate names: homeotic protein Chox-2.6; homeotic protein Chox-z

C;Species: Gallus gallus (chicken)

C;Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999

C;Accession: S10092

R;Sasaki, H.; Kuroiwa, A.

Nucleic Acids Res. 18, 184, 1990

A;Title: The nucleotide sequence of the cDNA encoding a chicken deformed family homeobox

A;Reference number: S10092; MUID:90174917; PMID:1968620

A;Accession: S10092

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-245 <SAS>

A;Cross-references: EMBL:X17612; NID:g63504; PIDN:CAA35614.1; PID:g63505

C;Function:

A;Description: control of embryonic development by tissue- and stage-specific regulation

C;Superfamily: homeotic protein Hox D4; homeobox homology

C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F;151-207/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 1; Length 245;

Best Local Similarity 90.0%; Pred. No. 1.2e-06;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20

DB 188 CLSERQIKIWFQNRMKWK 207

RESULT 17

T46446

hypothetical protein DKFPz434G0128.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 08-Sep-2000

C;Accession: T46446

R;Bloecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23032

A;Accession: T46446

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-246 <AAA>

A;Cross-references: EMBL:AL137449

A;Experimental source: adult testis; clone DKFPz434G0128

C;Genetics:

A;Note: DKFPz434G0128.1

C;Superfamily: homeotic protein Hox D4; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 83.9%; Score 99; DB 2; Length 246;

Best Local Similarity 90.0%; Pred. No. 1.2e-06;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20

DB 195 CLSERQIKIWFQNRMKWK 214

RESULT 18

A31757

homeotic protein Hox B4 - mouse

N;Alternate names: homeotic protein Hox 2.6

C;Species: Mus musculus (house mouse)

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 22-Jun-1999

C;Accession: A31757

R;Graham, A.; Papalopulu, N.; Lorimer, J.; McVey, J.H.; Tuddenham, E.G.D.; Krumlauf, R.

Genes Dev. 2, 1424-1438, 1988

A;Title: Characterization of a murine homeo box gene, Hox-2.6, related to the Drosophila

A;Reference number: A31757; MUID:89091992; PMID:2463210

A;Accession: A31757

A;Molecule type: DNA

A;Residues: 1-250 <GRA>

A;Cross-references: EMBL:M36654; NID:g193943; PIDN:AAA37848.1; PID:g193944

C;Genetics:

A;Gene: Hoxb-4

A;Map position: 11

A;Introns: 151/3

C;Function:

A;Description: control of embryonic development by tissue- and stage-specific regulation

C;Superfamily: homeotic protein Hox D4; homeobox homology

C;Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F;162-218/Domain: homeobox homology <HOX>

Query Match 83.9%; Score 99; DB 1; Length 250;

Best Local Similarity 90.0%; Pred. No. 1.2e-06;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMKWK 20

DB 199 CLSERQIKIWFQNRMKWK 218

RESULT 19

B60492

homeotic protein Hox B4 - human

N;Alternate names: homeotic protein Hox 2.6; homeotic protein Hox 2F

C;Species: Homo sapiens (man)

C;Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 22-Jun-1999

C;Accession: B60492; C37042; S15543

R;Reverali, F.A.; D'Esposito, M.; Acampora, D.; Bunone, G.; Negri, M.; Faiella, A.; Storti

Differentiation 45, 61-69, 1990

A;Title: Expression of HOX homeogenes in human neuroblastoma cell culture lines.

A;Reference number: A60492; MUID:91153613; PMID:1981366

A;Accession: B60492

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-251 <PEV>

R;Giampao, A.; Acampora, D.; Zappavigna, V.; Pannese, M.; D'Esposito, M.; Care, A.; Fa

Differentiation 40, 191-197, 1989

A;Title: Differential expression of human HOX-2 genes along the anterior-posterior axis

A;Reference number: A37042; MUID:89378558; PMID:2570724

A;Accession: C37042

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 160-227 <GIA>

A;Cross-references: GB:X16174; NID:G32376; PIDN:CAA34296.1; PID:g939889

R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Storti

Genome 31, 745-756, 1989

A;Title: Organization of human class I homeobox genes.

A;Reference number: S15036; MUID:90215256; PMID:2576652

A;Accession: S15543

A;Status: not compared with conceptual translation

RESULT 21
WJHU3E

homeotic protein Hox C4 - human
N;Alternate names: homeotic protein cpl9; homeotic protein cp8; homeotic protein Hox 3E
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S01030; S15545
R;Simone, A.; Pannese, M.; Acampora, D.; d'Esposito, M.; Boncinelli, E.
Nucleic Acids Res. 16, 5379-5390, 1988
A;Title: At least three human homeoboxes on chromosome 12 belong to the same transcript
A;Reference number: S01030; PMID:88262550; PMID:2898768
A;Accession: S01030
A;Molecule type: mRNA
A;Residues: 1-264 <SIM>
A;Cross-references: EMBL:X07495; NID:G32385; PIDN:CAA30376.1; PID:G32386
A;Note: The sequence from Fig. 4 is inconsistent with that from Fig. 3 in lacking 108-AL
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino, G.; Sci
Genome 31, 745-756, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; PMID:90215256; PMID:2576652
A;Accession: S15545
A;Molecule type: DNA
A;Residues: 156-221 <BON>
C;Genetics:
A;Gene: GDB:H0XC4
A;Cross-references: GDB:120672; OMIM:142974
A;Map position: 12q13.3-12q13.3
C;Function:
A;Description: control of embryonic development by tissue- and stage-specific regulation
C;Superfamily: homeotic protein Hox D4; homeobox homology
C;Keywords: alternative splicing; DNA binding; embryo; homeobox; nucleus; transcription
F;157-213/Domain: homeobox homology <Hox>

Query Match 83.9%; Score 99; DB 1; Length 264;
Best Local Similarity 90.0%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRRMKWK 20
| | | | | | | | | | | | | |
Db 194 CLSERQIKIWFQNRRMKWK 213

RESULT 22
S35219
homeotic protein Hox C4 - mouse
N;Alternate names: homeotic protein Hox 3.5; homeotic protein MAB87
C;Species: Mus musculus (house mouse)
C;Date: 10-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C;Accession: S35219; A49153; C41606; I49752
R;Goto, J.; Miyabayashi, T.; Wakamatsu, Y.; Takahashi, N.; Muramatsu, M.
Mol. Gen. Genet. 239, 41-48, 1993
A;Title: Organization and expression of mouse Hox3 cluster genes.
A;Reference number: S35219; PMID:93288004; PMID:8099712
A;Accession: S35219
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-264 <GT>
A;Cross-references: GB:S62287; NID:G385749; PIDN:AAB27153.1; PID:G385750; GB:D11328; NIH
A;Note: entry MUSHOX35A in GenBank release 103 duplicates GenBank entry SG2287 except for
E;Gaada, A.M.; Gaunt, S.J.; Azzawi, M.; Shimmel, S.M.; Pearce, J.; Sharpe, P.T.
Development 116, 497-506, 1992
A;Title: Sequence and embryonic expression of the murine Hox-3.5 gene.
A;Reference number: A49153; PMID:93161956; PMID:1363091
A;Accession: A49153
A>Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-79,'G','81'-95,'S','97'-264 <GEA>
A;Cross-references: GB:X69019; NID:G996183; PIDN:CXA48784.1; PID:G956184
A;Note: sequence extracted from NCBI backbone (NCBIN:124829, NCBI:P:124830)
R;Murtha, M.T.; Leckman, J.F.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 10711-10715, 1991
A;Title: Detection of homeobox genes in development and evolution.

Search completed: May 24, 2004, 17:26:11
Job time : 11.6216 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:17:30 ; Search time 7.2973 Seconds
(without alignments)
142.711 Million cell updates/sec

Title: US-09-977-349-3

Perfect score: 118

Sequence: 1 CSSCRQIKWQNRMRKWK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	83.9	49	1 HXA5 SHEEP	Q28599 ovis aries
2	99	83.9	76	1 HXC4 RAT	P18865 rattus norv
3	99	83.9	78	1 HXA5_SALSA	P09637 salmo salar
4	99	83.9	81	1 HXB4 BRARE	P09013 brachydanio
5	99	83.9	82	1 HXB5 CHICK	P14838 gallus gall
6	99	83.9	105	1 HXB4 BRARE	P22574 brachydanio
7	99	83.9	148	1 HXA5 AMBME	P50208 ambystoma m
8	99	83.9	225	1 HXA7 MORSA	Q9pww4 morone saxa
9	99	83.9	230	1 HXB5 XENLA	P09019 xenopus lae
10	99	83.9	233	1 HXA5 RAT	P52949 rattus norv
11	99	83.9	235	1 HXD4 CHICK	P17278 gallus gall
12	99	83.9	245	1 HXB4 CHICK	P14840 gallus gall
13	99	83.9	250	1 HXB4 MOUSE	P10284 mus musculu
14	99	83.9	250	1 HXD4 MOUSE	P10628 mus musculu
15	99	83.9	251	1 HXB4 FUGRU	O13074 fugu rubrip
16	99	83.9	251	1 HXB4 HUMAN	P17483 homo sapien
17	99	83.9	255	1 HXB4 HUMAN	P09016 homo sapien
18	99	83.9	264	1 HXC4 HUMAN	P09017 homo sapien
19	99	83.9	264	1 HXC4 MOUSE	Q08624 mus musculu
20	99	83.9	269	1 HXB5 HUMAN	P09067 homo sapien
21	99	83.9	269	1 HXB5 MOUSE	P09079 mus musculu
22	99	83.9	270	1 HXA5 MOUSE	P20719 homo sapien
23	99	83.9	270	1 HXA5 MOUSE	P09021 mus musculu
24	99	83.9	275	1 HXB5 BRARE	P09014 brachydanio
25	99	83.9	281	1 HXA5 MORSA	Q9pww3 morone saxa
26	98	83.1	67	1 HXA4 RAT	P09635 rattus norv
27	98	83.1	247	1 HXA4 HETER	Q9ia22 heterodontu
28	98	83.1	248	1 HXA4 MORSA	Q9pww2 morone saxa
29	98	83.1	285	1 HXA4 MOUSE	P06798 mus musculu
30	98	83.1	309	1 HXA4 CHICK	P17277 gallus gall
31	98	83.1	320	1 HXA4 HUMAN	Q00056 homo sapien
32	96	81.4	48	1 HXB6 XENLA	P31256 xenopus lae
33	96	81.4	71	1 HXA7 SHEEP	Q28600 ovis aries

34	96	81.4	71	1 HXC5 NOTVI	P31262 notophthalm
35	96	81.4	74	1 HWS0 APINE	P15860 apis mellif
36	96	81.4	75	1 HWSA SALSA	P09636 salmo salar
37	96	81.4	84	1 HXB6 CHICK	P14839 gallus gall
38	96	81.4	86	1 SCR APIME	P15859 apis mellif
39	96	81.4	87	1 HXC5 XENLA	P09020 xenopus lae
40	96	81.4	96	1 HXC6 BRARE	P15862 brachydanio
41	96	81.4	105	1 HXA7 RAT	P09634 rattus norv
42	96	81.4	112	1 HXB7 RAT	P18864 rattus norv
43	96	81.4	153	1 HXC6 SHEEP	P49925 ovis aries
44	96	81.4	208	1 HXA7 HETER	Q9ia25 heterodontu
45	96	81.4	209	1 HXA7 XENLA	P09071 xenopus lae
46	96	81.4	217	1 HXB7 BOVIN	Q9tt89 bos taurus
47	96	81.4	217	1 HXB7 HUMAN	P09629 homo sapien
48	96	81.4	217	1 HXB7 MOUSE	P09624 mus musculu
49	96	81.4	220	1 HXB7A XENLA	Q9ia771 xenopus lae
50	96	81.4	220	1 HB7B XENLA	P04476 xenopus lae
51	96	81.4	222	1 HXC5 HUMAN	Q00444 homo sapien
52	96	81.4	222	1 HXC5 MOUSE	P32043 mus musculu
53	96	81.4	224	1 HXB6 HUMAN	P17509 homo sapien
54	96	81.4	224	1 HXB6 MOUSE	P09023 mus musculu
55	96	81.4	228	1 HXB6 BRARE	P15861 brachydanio
56	96	81.4	229	1 HXA6 HETER	Q9ia24 heterodontu
57	96	81.4	229	1 HXA7 MOUSE	P02830 mus musculu
58	96	81.4	230	1 HXA7 HUMAN	P31268 homo sapien
59	96	81.4	232	1 HXA6 MOUSE	P09092 mus musculu
60	96	81.4	232	1 HXC5 BRARE	P09074 brachydanio
61	96	81.4	233	1 HXC6 NOTVI	P31267 homo sapien
62	96	81.4	234	1 HXC6 HUMAN	P14858 notophthalm
63	96	81.4	234	1 HXC6 XENLA	P02832 xenopus lae
64	96	81.4	235	1 HXC6 HUMAN	P09630 homo sapien
65	96	81.4	235	1 HXC6 MOUSE	P10629 mus musculu
66	96	81.4	242	1 HXA7 COTUA	P24061 coturix co
67	96	81.4	252	1 HXD5 HETER	Q9ia11 heterodontu
68	96	81.4	275	1 HXA5 HETER	Q9ia23 heterodontu
69	96	81.4	378	1 HXAN DROME	P02833 drosophila
70	96	81.4	394	1 HXAN DROSU	Q24645 drosophila
71	96	81.4	415	1 SCR DROME	P09077 drosophila
72	92	78.0	80	1 HXA4 LINS	P81192 lineus sang
73	92	78.0	87	1 HXA4 SHEEP	Q28598 ovis aries
74	92	78.0	93	1 HXB8 PIG	P09078 sus scrofa
75	92	78.0	232	1 HXB4 XENLA	P09070 xenopus lae
76	92	78.0	236	1 HXD4 BRARE	O57374 brachydanio
77	92	78.0	261	1 HXC4 ORVIA	O9pvs4 oryzias lat
78	92	78.0	590	1 HMDP DROME	P07548 drosophila
79	91	77.1	92	1 HXB8 RAT	P18863 rattus norv
80	91	77.1	108	1 HXC8 RAT	P18866 rattus norv
81	91	77.1	188	1 HXD8 CHICK	P23459 gallus gall
82	91	77.1	240	1 HXD8 HETER	Q9ia12 heterodontu
83	91	77.1	242	1 HXC8 HUMAN	P31273 homo sapien
84	91	77.1	242	1 HXC8 MOUSE	P09025 mus musculu
85	91	77.1	243	1 HXB8 HUMAN	P17481 homo sapien
86	91	77.1	243	1 HXB8 MOUSE	P09632 mus musculu
87	91	77.1	289	1 HXD8 MOUSE	P23463 mus musculu
88	91	77.1	290	1 HXD8 HUMAN	P13378 homo sapien
89	89	75.4	108	1 HMB3 TRIGR	P10178 tripneustes
90	87	73.7	60	1 HMA2 HELTR	P17138 helobdella
91	87	73.7	253	1 MOX1 MOUSE	P32442 mus musculu
92	87	73.7	254	1 MOX1 HUMAN	P50221 homo sapien
93	87	73.7	271	1 HMB XENLA	P14837 xenopus lae
94	87	73.7	283	1 IPF1 HUMAN	P52945 h insulinf p
95	87	73.7	283	1 IPF1 MESAU	P70118 mesocricetu
96	87	73.7	283	1 IPF1 RAT	P52947 rattus norv
97	87	73.7	284	1 IPF1 MOUSE	P52946 mus musculu
98	87	73.7	298	1 MOX2 XENLA	P39021 xenopus lae
99	87	73.7	303	1 MOX2 HUMAN	P50222 homo sapien
100	87	73.7	303	1 MOX2 MOUSE	P32443 mus musculu

ALIGNMENTS

RESULT 1

```

HXAS_SHEEP
ID HXAS_SHEEP STANDARD; PRT; 49 AA.
AC Q28599;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (Fragment).
GN HXAS OR HOXA-5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Roche P.J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC Also binds to its own promoter. Binds specifically to the motif:
CC 5'-CYNNATTA[GTG]Y-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U61978; AAB04754.1; -.
CC HSP; P02833; IHOM.
CC InterPro: IPR001356; Homeobox.
CC Pfam: PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC FT DNA_BIND <1 49 HOMEBOX.
CC FT NON_TER 49 49
CC SEQUENCE 49 AA; 6331 MW; 1EE702315E7C099B CRC64;

Query Match 83.9%; Score 99; DB 1; Length 49;
Best Local Similarity 90.0%; Pred. No. 1.6e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKWIFQNRMRKWK 20
DB 28 CLSERQIKWIFQNRMRKWK 47

RESULT 2
HX4 RAT
ID HX4 RAT STANDARD; PRT; 76 AA.
AC P18855;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-C4 (R3) (Fragment).
GN HOXC4 OR HOXC-4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Sprague-Dawley;
RX MEDLINE=89231502; PubMed=2907739;
RA Falzon M., Chung S.Y.;
RT "The expression of rat homeobox-containing genes is developmentally
RL regulated and tissue specific.";
RL Development 103:601-610(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system on the anterior-posterior axis.
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M37567; AAA41343.1; -.
CC PIR; C43559; C43559.
CC HSP; P02833; 9ANT.
CC InterPro: IPR001827; Antennapedia.
CC InterPro: IPR001356; Homeobox.
CC InterPro: IPR000047; HTH_lambrepresr.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEBOX.
CC PRINTS; PR00031; HTHREPRESSR.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC NON_TER 1 1
CC FT DNA_BIND 11 70 HOMEBOX.
CC FT SEQUENCE 76 AA; 9293 MW; 5235F665C0672385 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 76;
Best Local Similarity 90.0%; Pred. No. 2.5e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKWIFQNRMRKWK 20
DB 49 CLSERQIKWIFQNRMRKWK 68

RESULT 3
HX5 SALSA
ID HX5 SALSA STANDARD; PRT; 78 AA.
AC P09637;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (S12-B) (Fragment).
GN HOXA5.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88226009; PubMed=2897318;
RA Fjose A., Molven A., Eiken H.G.;
RT "Molecular cloning and characterization of homeo-box-containing genes
RT from Atlantic salmon.";
RL Gene 62:141-152(1988).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of

```

CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M18904; AAA49560.1; -.
CC PIR; I51342; I51342.
CC HGSP; P02833; 9ANT.

CC InterPro; IPR001827; Antennapedia.

CC InterPro; IPR001356; Homeobox.

CC Pfam; PF00046; homeobox; 1.

CC PRINTS; PR00024; HOMEBOX.

CC ProDom; PD000010; Homeobox; 1.

CC SMART; SM00389; HOX; 1.

CC PROSITE; PS00027; HOMEBOX_1; 1.

CC PROSITE; PS00032; ANTENNAPEIDIA; PARTIAL.

CC PROSITE; PS00071; HOMEBOX_2; 1.

CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.

CC NON_TER 1 1

CC DNA_BIND 1 60 HOMEBOX.

CC SEQUENCE 78 AA; 9489 MW; 828EBDDF78AC820 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 78;
Best Local Similarity 90.0%; Pred. No. 2.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKWIFQNRMRMKWK 20

Db 39 CLSERQIKWIFQNRMRMKWK 58

RESULT 4

HXSL BRARE

ID HXSL BRARE STANDARD; PRT; 81 AA.

AC P03013;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Homeobox protein Hox-B5 like (ZF-54) (Fragment).

GN HOXB5 OR ZF54 OR ZF-54.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.

OC NCBI_TaxID=7955;

CC [1]

CC SEQUENCE FROM N.A.

CC MEDLINE=89016617; PubMed=2902580;

CC Njolstad P.R., Molven A., Hordvik I., Apold J., Fjose A.;

CC "Primary structure, developmentally regulated expression and
CC potential duplication of the zebrafish homeobox gene ZF-21.";

CC Nucleic Acids Res. 16:9097-9113 (1988).

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

CC EMBL; X12803; CAA31291.1; -.

CC HSP; P02833; 1SAN.

CC ZFIN; ZDB-GENE-000823-6; hoxb5b.

CC InterPro; IPR001356; Homeobox.

CC Pfam; PF00046; homeobox; 1.

CC PRINTS; PR00024; HOMEBOX.

CC ProDom; PD000010; Homeobox; 1.

CC SMART; SM00389; HOX; 1.

CC PROSITE; PS00027; HOMEBOX_1; 1.

CC PROSITE; PS00071; HOMEBOX_2; 1.

CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.

CC NON_TER 1 1

CC DNA_BIND 1 65 HOMEBOX.

CC SEQUENCE 81 AA; 9977 MW; B7698AEFFEB3C6B4 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 81;
Best Local Similarity 90.0%; Pred. No. 2.7e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKWIFQNRMRMKWK 20

Db 44 CLSERQIKWIFQNRMRMKWK 63

RESULT 5

HXB5 CHICK

ID HXB5 CHICK STANDARD; PRT; 82 AA.

AC P14838;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Homeobox protein Hox-B5 (Glox-2.1) (Fragment).

GN HOXB5 OR GLOX-2.1.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

OC NCBI_TaxID=9031;

CC [1]

CC SEQUENCE FROM N.A.

CC TISSUE=Erythrocyte;

CC MEDLINE=90126373; PubMed=2575515;

CC Wedden S.E., Pang K., Bichele G.;

CC "Expression pattern of homeobox-containing genes during chick
CC embryogenesis.";

CC Development 105:639-650 (1989).

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X16846; CRA34743.1; -.

CC PIR; S08302; S08302.

CC HSP; P02833; 1SAN.

CC InterPro; IPR001827; Antennapedia.

CC InterPro; IPR001356; Homeobox.

CC Pfam; PF00046; homeobox; 1.

CC PRINTS; PR00024; HOMEBOX.

CC ProDom; PD000010; Homeobox; 1.

CC SMART; SM00389; HOX; 1.

CC PROSITE; PS00027; HOMEBOX_1; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
 DR PROSITE; PS50071; HOMEOBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT NON TER 1
 FT DNA_BIND 7 66 HOMEOBOX.
 SQ SEQUENCE 82 AA; 9877 MW; 53F70ACDC9FDEF8F CRC64;

Query Match 83.9%; Score 99; DB 1; Length 82;
 Best Local Similarity 90.0%; Pred. No. 2.7e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRKMKKK 20
 Db 45 CLSERQIKIWFQNRKMKKK 64

RESULT 6

HXB4 BRARE STANDARD; PRT; 105 AA.

AC F22574; O42369;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-B4 (ZF-13) (Fragment).
 GN HOXB4 OR HOXB4A OR HOXB4 OR ZF-13.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE OF 1-86 FROM N.A.
 RX MEDLINE=89211958; PubMed=2468579;
 RA Njoeistad P.R., Molven A., Eiken H.G., Fjose A.;
 RT "Structure and neural expression of a zebrafish homeobox sequence.";
 RL Gene 73:33-46(1988).
 RN [2]

SEQUENCE OF 44-105 FROM N.A.

RP Prince V.E., Moens C.B., Kimmel C.B., Ho R.K.;
 RA "Zebrafish hox genes: expression in the hindbrain region of wild-type
 and mutants of the segmentation gene, valentino.";
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; M24085; AAA56866.1; ALT INIT.
 CC EMBL; Y13946; CAA74284.1; -.
 CC PIR; JTO489; JTO489.
 CC HSSP; P02833; 9ANT.

DR ZFIN; ZDB-GENE-990415-105; hoxb4a.

DR InterPro; IPR001827; Antennapedia.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEOBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEOBOX_1; 1.

DR PROSITE; PS50071; HOMEOBOX_2; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;

Transcription regulation.

FT NON TER 1
 FT DNA_BIND 10 69 HOMEOBOX.
 SQ SEQUENCE 105 AA; 12262 MW; BUEFD84D909289F1 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 105;
 Best Local Similarity 90.0%; Pred. No. 3.5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRKMKKK 20
 Db 48 CLSERQIKIWFQNRKMKKK 67

RESULT 7

HXAS AMEME STANDARD; PRT; 148 AA.

AC P50208;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-A5 (Fragment).
 GN HOXA5.
 OS Ambystoma mexicanum (Axolotl).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;
 OC Ambystoma.
 OC NCBI_TaxID=8296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032352; PubMed=7557438;
 RA Gaur A.P., Lemanski L.F., Dube D.K.;
 RT "Identification and expression of a homologue of the murine HoxA5
 gene in the Mexican axolotl (Ambystoma mexicanum).";
 RL Gene 162:249-253(1995).
 CC -1- FUNCTION: Sequence-specific transcription factor which is part of
 a developmental regulatory system that provides cells with
 specific positional identities on the anterior-posterior axis (By
 similarity).

-1- SUBCELLULAR LOCATION: Nuclear.

-1- SIMILARITY: Belongs to the Antp homeobox family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; U19238; AAA91634.1; -.

CC PIR; PC4071; PC4071.

CC HSSP; P02833; LSAN.

CC TRANSFAC; T03305; -.

DR InterPro; IPR001827; Antennapedia.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00025; ANTENNAPEDIA.

DR PRINTS; PR00024; HOMEOBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEOBOX_1; 1.

DR PROSITE; PS00032; ANTENNAPEDIA; 1.

DR PROSITE; PS50071; HOMEOBOX_2; 1.

KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;

Transcription regulation.

FT NON TER 1

FT DNA_BIND 73 132 HOMEOBOX.

SQ SEQUENCE 148 AA; 16758 MW; C1893F0ED95F5086 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 148;

Best Local Similarity 90.0%; Pred. No. 5e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY      1  CSSCRQIKIWFQNRMMKWK 20
DB      111 CLSERQIKIWFQNRMMKWK 130

RESULT 8
HX57_MORSA
ID  HXA7_MORSA      STANDARD;      PRT;      225 AA.
AC  Q9PND4;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Homeobox protein Hox-A7.
GN  HXA7.
OS  Morone saxatilis (Striped bass).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC  Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC  Moronidae; Morone.
CX  NCBI_TaxID=34816;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99259633; PubMed=10327649;
RA  Snell E.A., Scemama J.L., Stollwag E.J.;
RT  "Genomic organization of the Hoxa4-Hoxa10 region from Morone
RT  saxatilis: implications for Hox gene evolution among vertebrates.";
RL  J. Exp. Zool. 285:41-49(1999).
CC  -!- FUNCTION: Sequence-specific transcription factor which is part of
CC  a developmental regulatory system that provides cells with
CC  specific positional identities on the anterior-posterior axis (By
CC  similarity).
CC  -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC  -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
DR  EMBL; AF089743; AAD46397.1; -.
DR  HSPF; P02833; 9ANT.
DR  InterPro; IPR001827; Antennapedia.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00025; ANTENNAPEDIA.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00032; ANTENNAPEDIA; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00711; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Transcription regulation.
FT  SITE 115 120      ANTP-TYPE HEXAPEPTIDE.
FT  SITE 126 185      HOMEBOX.
FT  DNA_BIND 126 185
SQ  SEQUENCE 225 AA; 25676 MW; C89FDEABA77FCBC CRC64;

Query Match      83.9%; Score 99; DB 1; Length 225;
Best Local Similarity 90.0%; Pred. No. 7.6e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CSSCRQIKIWFQNRMMKWK 20
DB      164 CLSERQIKIWFQNRMMKWK 183

RESULT 9
HXB5_XENLA
ID  HXB5_XENLA      STANDARD;      PRT;      230 AA.

```

```

AC  P09019;
DT  01-NOV-1988 (Rel. 09, Created)
DT  01-NOV-1988 (Rel. 09, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Homeobox protein Hox-B5 (XlHox-4) (Xhox-1B) (Fragment).
GN  HOB5 OR XLHOBX4.
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC  Xenopodinae; Xenopus.
CX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88157707; PubMed=2894634;
RA  Fritz A., de Robertis E.M.;
RT  "Xenopus homeobox-containing cDNAs expressed in early development.";
RL  Nucleic Acids Res. 16:1453-1469(1988).
CC  [2]
CC  SEQUENCE OF 155-214 FROM N.A.
CC  MEDLINE=86274626; PubMed=3015593;
CC  Harvey R.P., Tabin C.J., Melton D.A.;
CC  "Embryonic expression and nuclear localization of Xenopus homeobox
CC  (Xhox) gene products.";
CC  EMBL; J. 51237-1244(1986).
CC  -!- FUNCTION: Sequence-specific transcription factor which is part of
CC  a developmental regulatory system that provides cells with
CC  specific positional identities on the anterior-posterior axis.
CC  -!- SUBCELLULAR LOCATION: Nuclear.
CC  -!- DEVELOPMENTAL STAGE: Embryo.
CC  -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
DR  EMBL; X07104; CAA30125.1; -.
DR  EMBL; M26883; AAA49755.1; -.
DR  PIR; S00592; S00592.
DR  HSPF; P02833; 18AN.
DR  TRANSFAC; T03764; -.
DR  InterPro; IPR001827; Antennapedia.
DR  InterPro; IPR001356; Homeobox.
DR  Pfam; PF00046; homeobox; 1.
DR  PRINTS; PR00025; ANTENNAPEDIA.
DR  PRINTS; PR00024; HOMEBOX.
DR  ProDom; PD000010; Homeobox; 1.
DR  SMART; SM00389; HOX; 1.
DR  PROSITE; PS00027; HOMEBOX_1; 1.
DR  PROSITE; PS00032; ANTENNAPEDIA; 1.
DR  PROSITE; PS00711; HOMEBOX_2; 1.
KW  Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW  Transcription regulation.
FT  NON_TER 1
FT  SITE 137 142      ANTP-TYPE HEXAPEPTIDE.
FT  DNA_BIND 155 214      HOMEBOX.
SQ  SEQUENCE 230 AA; 25276 MW; AD040C030F85532B CRC64;

Query Match      83.9%; Score 99; DB 1; Length 230;
Best Local Similarity 90.0%; Pred. No. 7.8e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CSSCRQIKIWFQNRMMKWK 20
DB      193 CLSERQIKIWFQNRMMKWK 212

RESULT 10
HXA5_RAT
ID  HXA5_RAT      STANDARD;      PRT;      233 AA.

```



```

AC P52949;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5 (Hox-1.3) (Fragment).
OS HOXA5 OR HOXA-5 OR HOX-1.3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RX MEDLINE=94347374; PubMed=7915120;
RA Gorski D.H., Iepage D.F., Walsh K.;
RT "Cloning and sequence analysis of homeobox transcription factor cDNAs
RT with an incise-containing probe.";
RL BioTechniques 16:856-858(1994).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC Also binds to its own promoter. Binds specifically to the motif:
CC 5'-CCYNATTA[TG]Y-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; L03556; AAA67844.1; -.
DR PIR; I65197; I65197.
DR HSP; P02833; 1SAN.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT NON_TER 1 1
FT SITE 155 160 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 174 233 HOMEBOX.
FT NON_TER 233 233
FT SEQUENCE 233 AA; 25387 MW; 0937608EEDAF368C CRC64;

Query Match 83.9%; Score 99; DB 1; Length 233;
Best Local Similarity 90.0%; Pred. No. 7.9e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
DB 212 CLSERQIKIWFQNRMRKWK 231

RESULT 11
HXD4 CHICK
ID HXD4 CHICK STANDARD; PRT; 235 AA.
AC P17278;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-D4 (Chox-4).
OS HOXD4 OR HOXD-4 OR CHOX-4.

```

```

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90245562; PubMed=1970866;
RA Sasaki H., Yokoyama E., Kuroiwa A.;
RT "Specific DNA binding of the two chicken Deformed family homeodomain
RT proteins, Chox-1.4 and Chox-a-2";
RL Nucleic Acids Res. 18:1739-1747(1990).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- FUNCTION: BINDS TO SITES IN THE 5'-FLANKING SEQUENCE OF ITS CODING
CC REGION WITH VARIOUS AFFINITIES. THE CONSENSUS SEQUENCES OF THE
CC HIGH AND LOW AFFINITY BINDING SITES ARE TAATGA(C/G) AND CTAATTTT.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X52671; CAA36897.1; -.
DR EMBL; X52672; CAA36898.1; -.
DR HSP; P02833; 9ANT.
DR TRANSFAC; T01753; -.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 123 128 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 144 203 HOMEBOX.
FT DOMAIN 212 217 SER-RICH.
FT SEQUENCE 235 AA; 26662 MW; B7115D434033E4B5 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 235;
Best Local Similarity 90.0%; Pred. No. 8e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRMRKWK 20
DB 182 CLSERQIKIWFQNRMRKWK 201

RESULT 12
HXB4 CHICK
ID HXB4 CHICK STANDARD; PRT; 245 AA.
AC P14840;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-B4 (Chox-2).
OS HOXB4 OR CHOX-2.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

```

```

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174917; PubMed=1968620;
RA Sasaki H., Kuroiwa A.;
RT "The nucleotide sequence of the cDNA encoding a chicken Deformed
RL family homeobox gene, Chox-Z."
RL Nucleic Acids Res. 18:184-184(1990).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17612; CAA35614.1; -.
CC PIR; S10092; S10092.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T01726; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00024; ANTENNAPEDIA.
CC PRINTS; PR00025; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 129 134 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 150 209 HOMEBOX.
SQ SEQUENCE 245 AA; 27782 MW; 4949B200FEC44E91 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 245;
Best Local Similarity 90.0%; Pred. No. 8.3e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRKMKWK 20
DB 188 CLSERQIKIWFQNRKMKWK 207

RESULT 13
HXB4_MOUSE
ID HXB4_MOUSE STANDARD; PRT; 250 AA.
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-B4 (Hox-2.6).
OS Mus musculus (Mouse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89091992; PubMed=2463210;
RA Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,
RA Krumlauf R.;
RT "Characterization of a murine homeo box gene, Hox-2.6, related to the
RT Drosophila Deformed gene."
RL Genes Dev. 2:1424-1438(1988).

```

```

CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36654; AAA37848.1; -.
CC PIR; A31757; A31757.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T01728; -.
CC MGD; MGI:96185; Hoxb4.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 15 138 PRO-RICH; PART OF THE TRANSCRIPTIONAL
FT ACTIVATION DOMAIN.
FT DOMAIN 71 86 POLY-PRO.
FT SITE 140 145 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 161 220 HOMEBOX.
SQ SEQUENCE 250 AA; 27519 MW; D09D477A0E585BE6 CRC64;

Query Match 83.9%; Score 99; DB 1; Length 250;
Best Local Similarity 90.0%; Pred. No. 8.5e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRKMKWK 20
DB 199 CLSERQIKIWFQNRKMKWK 218

RESULT 14
HXD4_MOUSE
ID HXD4_MOUSE STANDARD; PRT; 250 AA.
AC P10628; P97451;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-D4 (Hox-4.2) (Hox-5.1).
GN HOXD4 OR HOXD-4 OR HOX-4.2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88263027; PubMed=2898782;
RA Featherstone M.S., Baron A., Gaunt S.J., Mattei M.-G., Duboule D.;
RT "Hox-5.1 defines a homeobox-containing gene locus on mouse chromosome
RT 2."
RL Proc. Natl. Acad. Sci. U.S.A. 85:4760-4764(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=94173687; PubMed=7907418;
RA Rambaldi I., Kovacs E.N., Featherstone M.S.;
RT "A proline-rich transcriptional activation domain in murine HOXD-4

```

```

RT (HOX-4.2).";
RL Nucleic Acids Res. 22:376-382(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Folberg A., Featherstone M.S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 175-198 FROM N.A.
RX MEDLINE=92212934; PubMed=1348361;
RA Nazatali A., Kim Y., Nirenberg M.;
RL "Hox-1.11 and Hox-4.9 homeobox genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:2883-2887(1992).
RX MEDLINE=97285126; PubMed=9140399;
RA Aparicio S.J., Hawker K., Cottage A., Mikawa Y., Zuo L., Venkatesh B.,
RL "Organizational evolution of the Fugu rubripes Hox clusters: evidence for
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03770; AAA20072.1; ALT_INIT.
CC EMBL; M87804; AAA37851.1; -.
CC EMBL; U77364; ABA41222.1; -.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T01752; -.
CC MGD; MGI:96208; Hoxd4.
CC
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS50071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC SITE 131 136 ANT-P-TYPE HEXAPEPTIDE.
CC FT DNA BIND 152 211 HOMEBOX.
CC FT DOMAIN 220 229 SER-RICH.
CC FT CONFLICT 143 143 A -> V (IN REF. 3).
CC FT SEQUENCE 250 AA; 27284 MW; 1057647C0A2665FF CRC64;
SQ
Query Match 83.9%; Score 99; DB 1; Length 250;
Best Local Similarity 90.0%; Pred. No. 8.5e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSSCRQIKIWFQNRMRKWK 20
DB 190 CLSERQIKIWFQNRMRKWK 209
RESULT 15
HXB4_FUGRU
ID HXB4_FUGRU STANDARD; PRT; 251 AA.
AC O13074;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-B4 (HOXB-4).
GN HOXB4.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95183530; PubMed=7878040;
RA Aparicio S., Morrison A., Gould A., Giltthorpe J., Chaudhuri C.,
RL Rigby P., Krumlauf R., Brenner S.;
RA "Detecting conserved regulatory elements with the model genome of the
RT Japanese puffer fish, Fugu rubripes.";
RN Proc. Natl. Acad. Sci. U.S.A. 92:1684-1688(1995).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=97285126; PubMed=9140399;
RA Aparicio S.J., Hawker K., Cottage A., Mikawa Y., Zuo L., Venkatesh B.,
RL Chen E., Krumlauf R., Brenner S.;
RA "Organizational evolution of the Fugu rubripes Hox clusters: evidence for
RT continuing evolution of vertebrate Hox complexes.";
RN Nat. Genet. 16:79-83(1997).
RX MEDLINE=97285126; PubMed=9140399;
RA Aparicio S.J., Hawker K., Cottage A., Mikawa Y., Zuo L., Venkatesh B.,
RL "Organizational evolution of the Fugu rubripes Hox clusters: evidence for
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U92575; AAC60205.1; -.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T03774; -.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX 1; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS50071; HOMEBOX 2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
CC Transcription regulation.
CC SITE 142 147 ANT-P-TYPE HEXAPEPTIDE.
CC FT DNA BIND 163 222 HOMEBOX.
CC FT SEQUENCE 251 AA; 28716 MW; 910D42998E68AFF1 CRC64;
SQ
Query Match 83.9%; Score 99; DB 1; Length 251;
Best Local Similarity 90.0%; Pred. No. 8.5e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSSCRQIKIWFQNRMRKWK 20
DB 201 CLSERQIKIWFQNRMRKWK 220
RESULT 16
HXB4_HUMAN
ID HXB4_HUMAN STANDARD; PRT; 251 AA.
AC P17483; Q9NTA0;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein Hox-B4 (Hox-2F) (Hox-2.6).
GN HOXB4 OR HOX2F.
OS Homo sapiens (Human).

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 (1)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91153613; PubMed=1981366;
 RA Peverali F.A., D'Esposito M., Acampora D., Runone G., Negri M.,
 RA Faiella A., Stornaiuolo A., Pannese M., Migliaccio E., Simeone A.,
 RA Valle G.D., Boncinelli E.;
 RT "Expression of HOX homeogenes in human neuroblastoma cell culture
 lines";
 RL Differentiation 45:61-69(1990).
 (2)
 RN SEQUENCE FROM N.A.
 RA Kidd K.K., Busygina V., DeMille M.M.C., Speed W.C., Ruggeri V.,
 RA Kidd J.R., Pakatis A.J.;
 RT "Overall linkage disequilibrium in 33 populations for highly
 RT informative multiseite haplotypes spanning the HOXB gene cluster";
 RL Am. J. Hum. Genet. 67:235-235(2000).
 (3)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20538492; PubMed=11085749;
 RA Giannola D.M., Shlomchik W.D., Jegathesan M., Liebowitz D.,
 RA Abrams C.S., Kadesch T., Dancis A., Emerson S.G.;
 RT "Hematopoietic expression of HOXB4 is regulated in normal and leukemic
 RT stem cells through transcriptional activation of the HOXB4 promoter by
 RT upstream stimulating factor (USF)-1 and USF-2";
 RL J. Exp. Med. 192:1479-1490(2000).
 (4)
 RN SEQUENCE FROM N.A.
 RT TISSUE=Uterus;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shervchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 (5)
 RN SEQUENCE OF 6-251 FROM N.A.
 RP TISSUE=Testis;
 RC Bloeker H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,
 RA Wiemann S.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 (6)
 RN SEQUENCE OF 160-227 FROM N.A.
 RP TISSUE=Placenta;
 RC MEDLINE=89378558; PubMed=2570724;
 RA Giampaolo A., Acampora D., Zappavigna V., Pannese M.,
 RA D'Esposito M., Care A., Faiella A., Stornaiuolo A., Russo G.,
 RA Simeone A., Boncinelli E., Peschle C.;
 RT "Differential expression of human HOX-2 genes along the anterior-
 RT posterior axis in embryonic central nervous system";
 RL Differentiation 40:191-197(1989).
 (7)
 RN SEQUENCE OF 162-227 FROM N.A.
 RP MEDLINE=90215256; PubMed=2576652;
 RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
 RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;

RT "Organization of human class I homeobox genes.";
 RL Genome 31:745-756(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DEVELOPMENTAL STAGE: Expressed in whole embryos and fetuses at
 CC 5-9 weeks from conception.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X16174; CAA34296.1; -;
 DR EMBL; AF287967; BAG31554.1; -;
 DR EMBL; AF307160; AAG45052.1; -;
 DR EMBL; BC049204; AAB49204.1; -;
 DR EMBL; AL137449; CAB70742.1; -;
 DR PIR; B60492; B60492.
 DR PIR; T46446; T46446.
 DR HSSP; P02833; 9ANT.
 DR TRANSFAC; T01727; -;
 DR Genew; HGNC:5115; HOXB4.
 DR MIM; 142965; -;
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; P:transcription factor activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00034; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.
 FT DOMAIN 15 139 PRO-RICH; PART OF THE TRANSCRIPTIONAL
 FT ACTIVATION DOMAIN.
 FT DOMAIN 71 87 POLY-PRO.
 FT SITE 141 146 ANTP-TYPE HEXAPEPTIDE.
 FT DNA_BIND 162 221 HOMEBOX.
 SQ SEQUENCE 251 AA; 27604 MW; 23EE1D0DDCCE2DB4 CRC64;
 Query Match 83.9%; Score 99; DB 1; Length 251;
 Best Local Similarity 90.0%; Pred. No. 8.5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CSSCRQIKIWFQNRMRKWK 20
 DB 200 CLSERQIKIWFQNRMRKWK 219
 RESULT 17
 HXD4 HUMAN STANDARD; PRT; 255 AA.
 ID HXD4 HUMAN
 AC P09016; Q9GAU0;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-D4 (Hox-4B) (Hox-5.1) (HHO.C13).
 GN HOXD4 OR HOXB4B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=90356367; PubMed=1975093;
 RA Cianetti L., di Cristofaro A., Zappavigna V., Bottero L.,
 Boccia G., Testa U., Russo G., Boncinelli E., Peschle C.,
 RT "Molecular mechanisms underlying the expression of the human HOX-5.1
 gene.";
 RL Nucleic Acids Res. 18:4361-4368(1990).
 RN [2]
 RP MEDLINE=22388257; PubMed=1247732;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP MEDLINE=154-219 FROM N.A.
 RX MEDLINE=90215256; PubMed=2576652;
 RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
 RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.,
 RT "Organization of human class I homeobox genes.";
 RL Genome 31:745-756(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X17360; CAA35237.1; -.
 CC EMBL; X04706; CAA28411.1; -.
 CC EMBL; BC016763; AAH16763.1; -.
 CC PIR; S10985; WJHU4B.
 CC HSPF; P02833; 9ANT.
 CC DR TRANSFAC; T00376; -.
 CC DR Genew; HGNC:5138; HOXD4.
 CC DR MIM; 142981; -.
 CC GO; GO:0005634; C:nucleus; NAS.

DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW transcription regulation.
 FT SITE 133 138 ANTP-TYPE HEXAPEPTIDE.
 FT DNA BIND 154 213 HOMEBOX.
 FT DOMAIN 222 234 POLY-SER.
 FT CONFLICT 123 123 S -> P (IN REF. 1).
 FT CONFLICT 142 142 V -> A (IN REF. 2).
 FT SEQUENCE 255 AA; 27885 MW; 39205343CB0142F CRC64;
 SQ
 Query Match 83.9%; Score 99; DB 1; Length 255;
 Best Local Similarity 90.0%; Pred. NO. 8.6e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CSSCROIKIWFQNRKWK 20
 DB 192 CLSERQIKIWFQNRKWK 211
 RESULT 18
 HXC4_HUMAN
 ID HXC4_HUMAN STANDARD; PRT; 264 AA.
 AC P09017;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-C4 (Hox-3E) (CP19).
 GN HoxC4 or Hox3E.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88252550; PubMed=2899768;
 RA Simeone A., Pannese M., Acampora D., D'Esposito M., Boncinelli E.,
 RT "At least three human homeoboxes on chromosome 12 belong to the same
 RT transcription unit.";
 RL Nucleic Acids Res. 16:5379-5390(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kosaki K., Kosaki R., Suzuki T., Yoshihashi H., Sasaki K., Matsuo N.,
 RT "A complete mutation analysis panel of human HOX genes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 156-221 FROM N.A.
 RX MEDLINE=90215256; PubMed=2576652;
 RA Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
 RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.,
 RT "Organization of human class I homeobox genes.";
 RL Genome 31:745-756(1989).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X07495; CAA30376.1; -;
 DR EMBL; AY014298; AAG42145.1; -;
 DR EMBL; AY014297; AAG42145.1; JOINED.
 DR F1R; S01030; WUHU3E.
 DR HSP; P02833; 9ANT.
 DR TRANSFAC; T03325; -;
 DR Genew; HGNC:5126; HOXC4.
 DR MIM; 142974; -;

DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0007275; P:development; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00027; HOMEBOX; 1.
 DR PROSITE; PS00071; HOMEBOX; 2; 1.
 DR Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.

FT DOMAIN 54 60 POLY-PRO.
 FT SITE 135 140 ANT-P-TYPE HEXAPEPTIDE.
 FT DNA_BIND 156 215 HOMEBOX.
 FT DOMAIN 183 186 POLY-ARG.
 SQ SEQUENCE 264 AA; 29781 MW; BDCD139955653373 CRC64;
 Query Match 83.9%; Score 99; DB 1; Length 264;
 Best Local Similarity 90.0%; Pred. No. 9e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSRCQIKIWFQNRMKWK 20
 Db 194 CLSERQIKIWFQNRMKWK 213
 RESULT 19
 HXC4 MOUSE STANDARD; PRT; 264 AA.
 ID HXC4 MOUSE STANDARD; PRT; 264 AA.
 AC Q08624;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Homeobox protein Hox-C4 (Hox-3.5).
 GN HOXC4 OR HOXC-4 OR HOXC-3.5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

CC STRAIN=BALB/c; TISSUE=Liver;
 CC MEDLINE=93288004; PubMed=8099712;
 RA Goto J., Miyabayashi T., Wakamatsu Y., Takahashi N., Muramatsu M.-A.;
 RT "organization and expression of mouse Hox3 cluster genes";
 RN Mol. Gen. Genet. 239:41-48(1993).
 RN [2]
 RP SEQUENCE FROM N.A.

CC TISSUE=Embryo;
 CC MEDLINE=93161956; PubMed=1363091;
 RA Gada A.M.C., Gaunt S.J., Azzawi M., Shimeld S.M., Pearce J.,
 RA Sharpe P.T.;
 RT "Sequence and embryonic expression of the murine Hox-3.5 gene";
 RN Development 116:497-506(1992).
 RN [3]
 RP SEQUENCE FROM N.A.

RP SEQUENCE OF 177-201 FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Spleen;
 RX MEDLINE=92073357; PubMed=1720547;
 RA Murtha W.T., Leckman J.F., Ruddle F.H.;
 RT "Detection of homeobox genes in development and evolution.";
 CC Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715(1991).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
 CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D11328; BAA01947.1; -;
 DR EMBL; S62287; AAB27153.1; -;
 DR EMBL; X69019; CAA48784.1; -;
 DR EMBL; M81660; AAG63313.1; -;
 DR F1R; S35219; S35219.
 DR HSP; P02833; 9ANT.
 DR TRANSFAC; T03340; -;
 DR MGD; MGI:96195; Hoxc4.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00027; HOMEBOX; 1; 1.
 DR PROSITE; PS00071; HOMEBOX; 2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 KW Transcription regulation.

FT DOMAIN 54 60 POLY-PRO.
 FT SITE 135 140 ANT-P-TYPE HEXAPEPTIDE (BY SIMILARITY).
 FT DNA_BIND 156 215 HOMEBOX (BY SIMILARITY).
 FT DOMAIN 183 186 POLY-ARG.
 FT CONFLICT 80 80 A -> G (IN REF. 2).
 FT CONFLICT 96 96 P -> S (IN REF. 2).
 SQ SEQUENCE 264 AA; 29865 MW; 35B0FA34B45BF30C CRC64;
 Query Match 83.9%; Score 99; DB 1; Length 264;
 Best Local Similarity 90.0%; Pred. No. 9e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSRCQIKIWFQNRMKWK 20
 Db 194 CLSERQIKIWFQNRMKWK 213
 RESULT 20
 HXB5 HUMAN STANDARD; PRT; 269 AA.
 ID HXB5 HUMAN STANDARD; PRT; 269 AA.
 AC P09067; P09069;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-B5 (Hox-2A) (Hox-C10) (HU-1).
 GN HOXB5 OR HOXA2A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=92385429; PubMed=1355360;
RA Galang C.K., Hauser C.A.;
RT "Cooperative DNA binding of the highly conserved human Hox 2.1
RT homeodomain gene product.";
RN New Biol. 4:558-568(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Kidd K.K., Busygina V., DeMille M.M.C., Speed W.C., Ruggeri V.,
RA Kidd J.R., Paketi A.J.;
RT "Overall linkage disequilibrium in 33 populations for highly
RT informative multiallelic haplotypes spanning the HOXB gene cluster.";
RN Am. J. Hum. Genet. 67:235-235(2000).
RN [3]
RP SEQUENCE OF 188-269 FROM N.A.
RX MEDLINE=86079489; PubMed=4075393;
RA Hauser C.A., Joyner A.L., Klein R.D., Learned T.K., Martin G.R.,
RA Tian R.;
RT "Expression of homologous homeo-box-containing genes in
RT differentiated human teratocarcinoma cells and mouse embryos.";
RN Cell 43:19-28(1985).
RN [4]
RP SEQUENCE OF 191-269 FROM N.A.
RX MEDLINE=85024858; PubMed=6091895;
RA Levine M., Rubin G.M., Tian R.;
RT "Human DNA sequences homologous to a protein coding region conserved
RT between homeotic genes of Drosophila.";
RN Cell 38:667-673(1984).
RN [5]
RP SEQUENCE OF 200-269 FROM N.A.
RX MEDLINE=86203637; PubMed=3453105;
RA Simeone A., Mavilio F., Bottero L., Giampaolo A., Russo G.,
RA Faiella A., Boncinelli E., Peschle C.;
RT "A human homeo box gene specifically expressed in spinal cord during
RT embryonic development.";
RN Nature 320:763-765(1986).
RN [6]
RP SEQUENCE OF 199-269 FROM N.A.
RX MEDLINE=90215286; PubMed=2576652;
RA Boncinelli E., Acampora D., Fannese M., D'Esposito M., Somma R.,
RA Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;
RT "Organization of human class I homeobox genes.";
RN Genome 31:745-756(1989).
RN [7]
RP SEQUENCE OF 217-269 FROM N.A.
RC TISSUE=Osteosarcoma;
RA Wayne M.M.Y.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: SPINAL CORD.
CC -!- DEVELOPMENTAL STAGE: Embryo.
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M92299; AAA52682.1; -;
CC EMBL; AF287967; AAG31553.1; -;
CC EMBL; X03794; CAA27420.1; -;
CC EMBL; K02572; AAA52681.1; ALT_INIT.
CC EMBL; M86726; AAB59430.1; -;
CC PIR; A24777; A24777.
CC PIR; A45578; A45578.
CC HSSP; P02833; 1SAN.
CC TRANSFAC; T01730; -;
CC Genew; HGNC:5116; HOXB5.

DR MIM: 142960; -;
DR GO: 0007345; P: embryogenesis and morphogenesis; TAS.
DR InterPro; IPR001827; Antennapedia.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 176 181 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 194 253 HOMEBOX.
SQ SEQUENCE 269 AA; 29434 MW; 58197F105DB0F8C4 CRC64;
Query Match 83.9%; Score 99; DB 1; Length 269;
Best Local Similarity 90.0%; Pred. No. 9.1e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CSSCRQIKIWFQNRMRMKWK 20
DB 232 CLSERQIKIWFQNRMRMKWK 251
RESULT 21
ID HXB5_MOUSE STANDARD; PRT; 269 AA.
AC P09079;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-B5 (Hox-2.1) (H24.1).
DE HOXB5 OR HOXB-5 OR HOX-2.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=88029099; PubMed=2889591;
RA Krumlauf R., Holland P.W.H., McVey J.H., Hogan B.L.M.;
RT "Developmental and spatial patterns of expression of the mouse
RT homeobox gene, Hox 2.1.";
RN Development 99:603-617(1987).
RN [2]
RP SEQUENCE OF 194-259 FROM N.A.
RX MEDLINE=86040438; PubMed=4058581;
RA Jackson I.J., Schofield P., Hogan B.L.M.;
RT "A mouse homeo box gene is expressed during embryogenesis and in
RT adult kidney.";
RN Nature 317:745-748(1985).
RN [3]
RP SEQUENCE OF 194-269 FROM N.A.
RX MEDLINE=86079489; PubMed=4075393;
RA Hauser C.A., Joyner A.L., Klein R.D., Learned T.K., Martin G.R.,
RA Tian R.;
RT "Expression of homologous homeo-box-containing genes in
RT differentiated human teratocarcinoma cells and mouse embryos.";
RN Cell 43:19-28(1985).
RN [4]
RP SEQUENCE OF 1-37 FROM N.A.
RA Arnold J.M.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
CC a developmental regulatory system that provides cells with
CC specific positional identities on the anterior-posterior axis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A BROAD SPECTRUM OF TISSUES.
CC -!- DEVELOPMENTAL STAGE: Embryo.

```

CC      -!- SIMILARITY: Belongs to the Antp homeobox family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M26283; AAA37842.1; -.
CC      EMBL; X03033; CAB57812.1; -.
CC      EMBL; U36070; AAA50294.1; -.
CC      PIR; A43551; A43551.
CC      PIR; B24777; B24777.
CC      HSSP; P02833; 1SAN.
CC      TRANSFAC; T01731; -.
CC      MGD; MGI:96186; Hoxb5.
CC      InterPro; IPR001827; Antennapedia.
CC      InterPro; IPR001356; Homeobox.
CC      Pfam; PF00046; homeobox; 1.
CC      PRINTS; PR00025; ANTENNAPEDIA.
CC      PRINTS; PR00024; HOMEBOX.
CC      ProDom; PD000010; Homeobox; 1.
CC      SMART; SM00389; HOX; 1.
CC      PROSITE; PS00027; HOMEBOX 1; 1.
CC      PROSITE; PS00032; ANTENNAPEDIA; 1.
CC      PROSITE; PS00071; HOMEBOX 2; 1.
CC      Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW      Transcription regulation.
FT      SITE 176 181
FT      DNA BIND 194 253
FT      SEQUENCE 269 AA; 29464 MW; 38576350CC95B95D CRC64;
SQ
Query Match 83.9%; Score 99; DB 1; Length 269;
Best Local Similarity 90.0%; Pred. No. 9.1e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CSSCRQIKWFFQRRMKWK 20
DB      232 CLSERQIKWFFQRRMKWK 251

RESULT 22
HXAS HUMAN
ID      HXAS HUMAN STANDARD; PRT; 270 AA.
AC      P20719; O43367;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Homeobox protein Hox-A5 (Hox-1C).
GN      HOXA5 OR HOX1C.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89313782; PubMed=2568583;
RA      Tournier-Lasserre E., Odenwald W.F., Garbern J., Trojanowski J.,
RA      Lazzarini R.A.;
RT      Remarkable intron and exon sequence conservation in human and mouse
RL      homeobox Hox 1.3 genes.
RL      Mol. Cell. Biol. 9:2273-2278 (1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Bradshaw H., Hinds K., Keppler D.;
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 195-260 FROM N.A.
RX      MEDLINE=90215256; PubMed=2576652;
RA      Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,
RA      Gaudino G., Stornaiuolo A., Cafiero M., Faiella A., Simeone A.;

```

```

RT      "Organization of human class I homeobox genes.";
RL      Genome 31:745-756 (1989).
CC      -!- FUNCTION: Sequence-specific transcription factor which is part of
CC      a developmental regulatory system that provides cells with
CC      specific positional identities on the anterior-posterior axis.
CC      Also binds to its own promoter. Binds specifically to the motif:
CC      5'-CYNAYTA[TG]Y-3'.
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- DEVELOPMENTAL STAGE: Expressed during embryogenesis and in adult
CC      kidney.
CC      -!- SIMILARITY: Belongs to the Antp homeobox family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M26679; AAA58663.1; -.
CC      EMBL; AC004080; -; NOT_ANNOTATED_CDS.
CC      PIR; A24777; A24777.
CC      PIR; A32799; WJHULC.
CC      HSSP; P02833; 1SAN.
CC      TRANSFAC; T01702; -.
CC      Genew; HGNC:5106; HOXA5.
CC      MIM; 142952; -.
CC      GO; GO:0005634; C:nucleus; NAS.
CC      GO; GO:0003700; F:transcription factor activity; NAS.
CC      GO; GO:0007275; P:development; NAS.
CC      GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC      InterPro; IPR001827; Antennapedia.
CC      InterPro; IPR001356; Homeobox.
CC      Pfam; PF00046; homeobox; 1.
CC      PRINTS; PR00025; ANTENNAPEDIA.
CC      PRINTS; PR00024; HOMEBOX.
CC      ProDom; PD000010; Homeobox; 1.
CC      SMART; SM00389; HOX; 1.
CC      PROSITE; PS00027; HOMEBOX 1; 1.
CC      PROSITE; PS00032; ANTENNAPEDIA; 1.
CC      PROSITE; PS00071; HOMEBOX 2; 1.
CC      Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW      Transcription regulation.
FT      SITE 176 181
FT      DNA BIND 195 254
FT      CONFLICT 114 114
FT      SEQUENCE 270 AA; 29359 MW; 6583BF22562BC9AC CRC64;
SQ
Query Match 83.9%; Score 99; DB 1; Length 270;
Best Local Similarity 90.0%; Pred. No. 9.2e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CSSCRQIKWFFQRRMKWK 20
DB      233 CLSERQIKWFFQRRMKWK 252

RESULT 23
HXAS MOUSE
ID      HXAS MOUSE STANDARD; PRT; 270 AA.
AC      P09021;
DT      01-NOV-1988 (Rel. 09, Created)
DT      01-NOV-1988 (Rel. 09, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Homeobox protein Hox-A5 (Hox-1.3) (M2).
GN      HOXA5 OR HOXA-5 OR HOX-1.3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.

```


RX MEDLINE=8056292; PubMed=2890554;
 RA Odenwald W.F., Taylor C.F., Palmer-Hill F.J., Friedrich V. Jr.,
 RA Tani M., Lazzarini R.A.;
 RT "Expression of a homeo domain protein in noncontact-inhibited
 RT cultured cells and postmitotic neurons.";
 RL Genes Dev. 1:482-496(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88328807; PubMed=2901335;
 RA Fibi M., Zink B., Kessel M., Colberg-Poley A.M., Labelt S.,
 RA Lehrach H., Gruss P.;
 RT "Coding sequence and expression of the homeobox gene Hox 1.3.";
 RL Development 102:349-359(1988).
 RN [3]
 RP DNA-BINDING SPECIFICITY.
 RX MEDLINE=89232713; PubMed=2565857;
 RA Odenwald W.F., Garbern J., Arnheiter H., Tournier-Lasserre E.,
 RA Lazzarini R.A.;
 RT "The Hox-1.3 homeo box protein is a sequence-specific DNA-binding
 RT phosphoprotein.";
 RL Genes Dev. 3:158-172(1989).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=96205869; PubMed=8635464;
 RA Zhao J.J., Lazzarini R.A., Pick L.;
 RT "Functional dissection of the mouse Hox-a5 gene.";
 RL EMBO J. 15:1313-1322(1996).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC Also binds to its own promoter. Binds specifically to the motif:
 CC 5'-CYNNATTA[GTG]-3'.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryogenesis and in adult
 CC kidney.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M36604; AAA37838.1; -;
 CC EMBL; Y00208; CAA68364.1; -;
 CC EMBL; X16840; CAA34738.1; -;
 CC EMBL; M28021; AAA37837.1; -;
 CC PIR; S07812; WJMS13.
 CC HSSP; P02833; ISAN.
 CC TRANSFAC; T00377; -;
 CC MGD; MGI:96177; Hoxa5.
 CC InterPro; IPR001827; Antennapedia.
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00025; ANTENNAPEDIA.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX 1; 1.
 CC PROSITE; PS00032; ANTENNAPEDIA; 1.
 CC PROSITE; PS50071; HOMEBOX 2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC Transcription regulation.
 KW SITE 176 181 ANT-P-TYPE HEXAPEPTIDE.
 FT DNA BIND 195 254 HOMEBOX
 FT SEQUENCE 270 AA; 29237 MW; DC4BDA8FE627665 CRC64;
 SQ
 Query Match 83.9%; Score 99; DB 1; Length 270;
 Best Local Similarity 90.0%; Pred. No. 9.2e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRKMKWK 20
 Db 233 CLSERQIKIWFQNRKMKWK 252
 RESULT 24
 HXB5 BRARE
 ID HXB5 BRARE STANDARD; PRT; 275 AA.
 AC P09014;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-B5 (ZF-21).
 GN HXB5 OR HOXB5A OR Hox-B5 OR ZF-21.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89016617; PubMed=2902580;
 RA Njolstad P.R., Molven A., Hordvik I., Apold J., Fjose A.;
 RT "Primary structure, developmentally regulated expression and
 RT potential duplication of the zebrafish homeobox gene ZF-21.";
 RL Nucleic Acids Res. 16:9097-9113(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93250038; PubMed=8097929;
 RA Molven A., Hordvik I., Njolstad P.R.;
 RT "Sequence analysis of the zebrafish hox-B5/B6 region.";
 RL Biochim. Biophys. Acta 1173:102-106(1993).
 RN [3]
 RP SEQUENCE OF 195-275 FROM N.A.
 RX MEDLINE=88167192; PubMed=2895022;
 RA Njolstad P.R., Molven A., Fjose A.;
 RT "A zebrafish homologue of the murine Hox-2.1 gene.";
 RL FEBS Lett. 230:125-30(1988).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the Antp homeobox family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X12802; CAA31290.1; -;
 CC EMBL; X07381; CAA30293.1; -;
 CC EMBL; X68248; CAA48320.1; -;
 CC PIR; S03671; WJZFX2.
 CC HSSP; P02833; ISAN.
 CC TRANSFAC; T03663; -;
 CC ZFIN; ZDB-GENE-980526-70; hoxb5a.
 CC InterPro; IPR001827; Antennapedia.
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00025; ANTENNAPEDIA.
 CC PRINTS; PR00024; HOMEBOX.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX 1; 1.
 CC PROSITE; PS00032; ANTENNAPEDIA; 1.
 CC PROSITE; PS50071; HOMEBOX 2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
 CC Transcription regulation.
 KW SITE 182 187 ANT-P-TYPE HEXAPEPTIDE.
 FT DNA BIND 200 259 HOMEBOX.

SQ SEQUENCE 275 AA; 30755 MW; F9CCD682138C626A CRC64;
Query Match 83.9%; Score 99; DB 1; Length 275;
Best Local Similarity 90.0%; Pred. No. 9.3e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRKMKWK 20
Db 238 CLSERQIKIWFQNRKMKWK 257

Db 243 CLSERQIKIWFQNRKMKWK 262

Search completed: May 24, 2004, 17:25:16
Job time : 8.2973 secs

RESULT 25

HXA5 MORSA
ID HXA5 MORSA STANDARD; PRT; 281 AA.
AC Q9PMD3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-A5.
GN HOXA5.
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Morone.
OX NCBI_TaxID=34816;
RN [1]_TaxID=34816;
RP SEQUENCE FROM N.A.
RX MEDLINE=99259633; PubMed=10327649;
RA Snell E.A., Scemama J.L., Stellwag E.J.;
RT "Genomic organization of the Hox4-Hoxa10 region from Morone
saxatilis: implications for Hox gene evolution among vertebrates.";
RL J. Exp. Zool. 285:41-49 (1999).
CC -!- FUNCTION: Sequence-specific transcription factor which is part of
a developmental regulatory system that provides cells with
specific positional identities on the anterior-posterior axis (By
similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the Antp homeobox family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF089743; AAD46398.1; -.
CC HSSP; P02833; 1SAN.
CC InterPro; IPR001827; Antennapedia.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox; 1.
CC PRINTS; PR00025; ANTENNAPEDIA.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD000010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00032; ANTENNAPEDIA; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT SITE 187 192 ANT-P-TYPE HEXAPEPTIDE.
FT DNA_BIND 205 264 HOMEBOX.
FT DOMAIN 232 235 POLY-ARG.
FT DOMAIN 275 278 POLY-GLY.
SQ SEQUENCE 281 AA; 31123 MW; 5726D13D47C8AE35 CRC64;

QY 1 CSSCROIKIWFQNRKMKWK 20
Query Match 83.9%; Score 99; DB 1; Length 281;
Best Local Similarity 90.0%; Pred. No. 9.5e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:14:23 ; Search time 34.5946 Seconds

(without alignments)
182.409 Million cell updates/sec

Title: US-09-977-349-3

Perfect score: 118

Sequence: 1 CSSCRQIKIWFQNRMRMKWK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL 25:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_rvirus:**
- 16: sp_bacteriopl:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	83.9	43	13	057359
2	99	83.9	57	13	Q9PVR8
3	99	83.9	60	5	077143
4	99	83.9	60	5	077139
5	99	83.9	60	13	Q8QGL8
6	99	83.9	60	13	Q8QGL7
7	99	83.9	61	5	Q27910
8	99	83.9	69	5	Q9BNF7
9	99	83.9	71	13	Q9PVS3
10	99	83.9	71	13	Q9PVS1
11	99	83.9	77	5	044260
12	99	83.9	81	13	Q9PVR7
13	99	83.9	88	5	Q967W7
14	99	83.9	89	5	Q967V4
15	99	83.9	92	13	057377
16	99	83.9	94	5	Q967V6

17	99	83.9	107	5	Q8WRM8
18	99	83.9	109	5	Q94575
19	99	83.9	109	5	Q05011
20	99	83.9	115	5	044258
21	99	83.9	115	13	Q801B5
22	99	83.9	146	5	Q05009
23	99	83.9	150	11	Q61679
24	99	83.9	225	13	Q801C0
25	99	83.9	227	13	Q9YGT6
26	99	83.9	227	13	Q9PVM5
27	99	83.9	230	11	Q9CZV0
28	99	83.9	235	13	Q9W7Q8
29	99	83.9	244	13	Q8AWY6
30	99	83.9	244	13	Q801C1
31	99	83.9	246	13	Q9PML9
32	99	83.9	263	5	076843
33	99	83.9	264	4	Q86TF7
34	99	83.9	270	4	Q96CY6
35	99	83.9	270	11	Q91VQ5
36	99	83.9	275	13	Q8AWY7
37	99	83.9	276	13	Q8JH53
38	99	83.9	276	13	Q9PVM0
39	99	83.9	277	5	Q9NGR4
40	99	83.9	303	13	Q9PML2
41	99	83.9	381	13	Q90WV0
42	98	83.1	39	6	Q8XT68
43	98	83.1	60	13	Q8QGL9
44	98	83.1	81	13	Q9PVS5
45	98	83.1	85	11	Q63255
46	98	83.1	85	11	Q61681
47	98	83.1	89	13	Q42371
48	98	83.1	106	13	Q801D2
49	98	83.1	193	5	Q817D0
50	98	83.1	244	13	Q9PML5
51	98	83.1	285	11	Q8BPB6
52	96	81.4	33	5	Q86FU0
53	96	81.4	39	13	057368
54	96	81.4	42	11	Q80WH6
55	96	81.4	46	13	Q9PVR9
56	96	81.4	51	5	Q23743
57	96	81.4	51	5	Q26407
58	96	81.4	58	5	Q25208
59	96	81.4	58	13	057362
60	96	81.4	59	5	Q8WRM9
61	96	81.4	59	13	Q9PVT5
62	96	81.4	60	11	Q80WH7
63	96	81.4	60	11	Q80WH4
64	96	81.4	60	13	Q8QGL5
65	96	81.4	60	13	Q8QGL3
66	96	81.4	60	13	Q8QGL6
67	96	81.4	60	13	Q8QGL2
68	96	81.4	63	5	077138
69	96	81.4	63	5	Q8WRM1
70	96	81.4	66	13	057356
71	96	81.4	70	5	Q967W5
72	96	81.4	70	13	Q801B4
73	96	81.4	73	5	Q9Y186
74	96	81.4	75	5	Q25209
75	96	81.4	75	13	Q9PVR6
76	96	81.4	76	5	044257
77	96	81.4	76	5	Q86NE1
78	96	81.4	77	5	Q9Y187
79	96	81.4	77	5	Q9U924
80	96	81.4	79	5	Q967V2
81	96	81.4	79	5	Q9U9T9
82	96	81.4	80	5	Q05008
83	96	81.4	81	5	Q9BN27
84	96	81.4	81	5	Q17142
85	96	81.4	82	5	Q24758
86	96	81.4	82	13	Q9PVS0
87	96	81.4	83	13	Q801B8
88	96	81.4	83	5	Q17140
89	96	81.4	85	13	Q801B9

Q8WRM8 lithobius a
Q94575 heliocidari
Q05011 artemia san
044258 ethmostigm
Q801B5 latimeria m
Q05009 artemia san
Q61679 mus musculu
Q801C0 latimeria m
Q9YGT6 brachydano
Q9PVM5 brachydano
Q9CZV0 mus musculu
Q9W7Q8 paralichthy
Q8AWY6 brachydano
Q801C1 latimeria m
Q9PML9 brachydano
076843 cupiemius
Q86TF7 homo sapien
Q96CY6 homo sapien
Q91VQ5 mus musculu
Q8AWY7 brachydano
Q8JH53 brachydano
Q9PVM0 brachydano
Q9NGR4 herdmania c
Q9PML2 petromyzon
Q90WV0 petromyzon
Q8XT68 canis famil
Q8QGL9 petromyzon
Q9PVS5 oryzias lat
Q63255 rattus norv
Q61681 mus musculu
Q42371 brachydano
Q801D2 latimeria m
Q817D0 ciona intes
Q9PML5 brachydano
Q8BPB6 mus musculu
Q86FU0 drosophila
Q57368 brachydano
Q80WH6 rattus sp.
Q9PVR9 oryzias lat
Q23743 ctenodrilus
Q26407 ctenodrilus
Q25208 junonia coe
Q57362 brachydano
Q8WRM9 lithobius a
Q9PVS5 oryzias lat
Q80WH7 rattus sp.
Q80WH4 rattus sp.
Q8QGL5 petromyzon
Q8QGL3 petromyzon
Q8QGL6 petromyzon
Q8QGL2 petromyzon
077138 archegozete
Q8WRM1 lithobius a
Q57356 brachydano
Q967W5 folsomia ca
Q801B4 latimeria m
Q9Y186 priapulid c
Q25209 junonia coe
Q9PVR6 oryzias lat
044257 ethmostigm
Q86NB1 callanus hel
Q9Y187 priapulid c
Q9U924 lingula ung
Q967V2 lithobius f
Q9U9T9 nereis vire
Q05008 artemia san
Q9BN27 porcellio s
Q17142 brachiosteo
Q24758 drosophila
Q9PVS0 oryzias lat
Q801B8 latimeria m
Q17140 brachiosteo
Q801B9 latimeria m

90 Q9w7p8 paralichthy
91 Q8qiv4 lampetra fl
92 Q8i7c9 ciona intes
93 Q8wgr9 eupyrima sc
94 Q8mx73 sacculina c
95 Q9gtm5 tribolium c
96 Q9ped5 gallus gall
97 Q62551 lineus sang
98 Q26498 schistocerc
99 Q8jzw2 mus musculu
100 Q44368 acanthokara

ALIGNMENTS

```

RESULT 1
O57359 PRELIMINARY; PRT; 43 AA.
AC O57359;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hoxa5 protein (Fragment).
GN HOXB5 OR HOXA5
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Prince V.E., Joly L., Ekker M., Ho R.K.;
RT "Zebrafish hox genes: genomic organization and modified colinear
RT expression patterns in the trunk."
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; Y14526; CAA74861.1; -.
DR ZFIN; ZDB-GENE-000823-6; hox5b.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 43 AA; 5050 MW; 53034C37F3DFA596 CRC64;

Query Match 83.9%; Score 99; DB 13; Length 43;
Best Local Similarity 90.0%; Pred. No. 5.4e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKIWFQNRMKWK 20
Db 6 CLSERQIKIWFQNRMKWK 25

RESULT 2
O9PVR8 PRELIMINARY; PRT; 57 AA.
AC O9PVR8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HOXA5A (Fragment).
GN HOXA5A.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo S., Naruse K., Shima A.;
RT "Hox genes of the medakafish Oryzias latipes."
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AB026961; BAA86244.1; -.
DR HSSP; P02833; IHOW.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 57 AA; 6891 MW; 54A6430320F68C04 CRC64;

Query Match 83.9%; Score 99; DB 13; Length 57;
Best Local Similarity 90.0%; Pred. No. 7.1e-08;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCRQIKIWFQNRMKWK 20
Db 19 CLSERQIKIWFQNRMKWK 38

RESULT 3
O77143 PRELIMINARY; PRT; 60 AA.
AC O77143;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sex combs reduced (Fragment).
GN SCR.
OS Archegozetes longisetosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Oribatida; Desmonomata;
OC Thysanothroniidae; Thysanothroniidae; Archegozetes.
OX NCBI_TaxID=66560;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98393703; PubMed=9724762;
RA Telford M.J., Thomas R.H.;
RT "Expression of homeobox genes shows chelicerate arthropods retain
RT their deutocerebral segment."
RL Proc. Natl. Acad. Sci. U.S.A. 95:10671-10675(1998).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF071407; AAC35936.1; -.
DR HSSP; P02833; ISAN.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
FT NON_TER 60

```

RN SEQUENCE FROM N.A.
 RP Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 RL marinus";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR ENBL; AF410911; AAM19469.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR000047; HTH_lambdarepress.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESS.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 60 60
 SQ SEQUENCE 60 AA; 7704 MW; 08EBB642C24DEC6E CRC64;

 Query Match 83.9%; Score 99; DB 13; Length 60;
 Best Local Similarity 90.0%; Pred. No. 7.4e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps

 QY 1 CSSCROIKIWFQNRMRMKKK 20
 DB 39 CLSEROIKIWFQNRMRMKKK 58

 RESULT 6
 C8QGL7 PRELIMINARY; PRT; 60 AA.
 ID C8QGL7
 AC C8QGL7;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DT Hox5 homeobox (Fragment).
 OS Petromyzon marinus (Sea lamprey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RL [1]
 RN SEQUENCE FROM N.A.
 RP Irvine S.Q., Carr J.L., Bailey W.J., Kawasaki K., Shimizu N.,
 RA Amemiya C.T., Ruddle F.H.;
 RT "Genomic analysis of Hox clusters in the sea lamprey Petromyzon
 RL marinus";
 RL J. Exp. Zool. 0:0-0(2002).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR ENBL; AF410912; AAM19470.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT NON_TER 1 1
 FT NON_TER 60 60
 SQ SEQUENCE 60 AA; 7635 MW; E1F1209C5ACDF866 CRC64;

 Query Match 83.9%; Score 99; DB 13; Length 60;
 Best Local Similarity 90.0%; Pred. No. 7.4e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 CSSCROIKIWFQNRMKKK 20
 Db 39 CLSERQIKIWFQNRMKKK 58

RESULT 7

Q27910 Q27910 PRELIMINARY; PRT; 61 AA.
 AC Q27910; Q94684;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Homeodomain protein PHOX5 (Fragment).
 GN PHOX5
 OS Polyandrocarpa misakiensis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Styelidae; Polyandrocarpa.
 OX NCBI_TaxID=7723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USHIMADO;
 RA Fujiwara S., Ohashi M., Kawamura K.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-25 FROM N.A.
 RC STRAIN=USHIMADO;
 RA Fujiwara S., Kawamura K.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; D86851; BAA13173.1; -.
 DR EMBL; D88277; BAA13573.1; -.
 DR HSP; P02833; IHOM.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRODOM; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 61 AA; 7537 MW; 14C09CGBAF112C33 CRC64;

Query Match 83.9%; Score 99; DB 5; Length 61;
 Best Local Similarity 90.0%; Pred. No. 7.5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 CSSCROIKIWFQNRMKKK 20
 Db 19 CLSERQIKIWFQNRMKKK 38

RESULT 8

Q9BMF7 Q9BMF7 PRELIMINARY; PRT; 69 AA.
 AC Q9BMF7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Homeobox protein Hox4 (Fragment).
 GN Hox4.
 OS Haliotis asinina.
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Vetigastropoda; Haliotoidea; Haliotidae; Haliotis.
 OX NCBI_TaxID=109174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USHIMADO;
 RA Fujiwara S., Ohashi M., Kawamura K.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-25 FROM N.A.
 RC STRAIN=USHIMADO;
 RA Fujiwara S., Kawamura K.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; D86851; BAA13173.1; -.
 DR EMBL; D88277; BAA13573.1; -.
 DR HSP; P02833; IHOM.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRODOM; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 61 AA; 7537 MW; 14C09CGBAF112C33 CRC64;

Query Match 83.9%; Score 99; DB 5; Length 61;
 Best Local Similarity 90.0%; Pred. No. 7.5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

SEQUENCE FROM N.A.
 RA Hinman V.F., Degnan B.M.;
 RT "Expression of anterior Hox genes during larval development of the
 RT gastropod Haliotis asinina";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF328863; AAK11240.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRODOM; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 69 AA; 8362 MW; 53C227503129AF06 CRC64;

Query Match 83.9%; Score 99; DB 5; Length 69;
 Best Local Similarity 90.0%; Pred. No. 8.5e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 CSSCROIKIWFQNRMKKK 20
 Db 17 CLSERQIKIWFQNRMKKK 36

RESULT 9

Q9PVS3 Q9PVS3 PRELIMINARY; PRT; 71 AA.
 AC Q9PVS3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE HOXB4A (Fragment).
 GN HOXB4A.
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kondo S., Naruse K., Shima A.;
 RT "Hox genes of the medakafish Oryzias latipes";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AB026956; BAA86239.1; -.
 DR HSP; P02833; IHOM.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRODOM; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS50071; HOMEBOX 2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 SQ SEQUENCE 71 AA; 8380 MW; 8793605DA3CD4D63 CRC64;

Query Match 83.9%; Score 99; DB 13; Length 71;
 Best Local Similarity 90.0%; Pred. No. 8.7e-08;
 Matches 18; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

QY 1 CSSCROIKIWFQNRMKKK 20
 Db 17 CLSERQIKIWFQNRMKKK 36

Db 19 CLSERQIKIWFQNRKMKWK 38

RESULT 10

Q9PVS1 PRELIMINARY; PRT; 71 AA.

AC Q9PVS1

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE HOXD4A (Fragment)

GN HOXD4A

OS Oryzias latipes (Medaka fish) (Japanese ricefish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

OX NCBI_TaxID=8090;

RN [1]

RP SEQUENCE FROM N.A.

RA Kondo S., Naruse K., Shima A.;

RT "Hox genes of the medakafish Oryzias latipes.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AB026598; BAA86241.1; -.

DR HSSP; P02833; 1HOM.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003700; P:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox.1.

DR PRINTS; PR00024; HOMEBOX.

DR SMART; SM00010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS00071; HOMEBOX 2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON_TER 1

FT SEQUENCE 71 AA; 8623 MW; 5D6E3DF055216227 CRC64;

Query Match 83.9%; Score 99; DB 13; Length 71;

Best Local Similarity 90.0%; Pred. No. 8.7e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKMKWK 20

Db 19 CLSERQIKIWFQNRKMKWK 38

RESULT 11

O44260 PRELIMINARY; PRT; 77 AA.

AC O44260

DT 01-JUN-1998 (TRENBLrel. 06, Created)

DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE Sex combs reduced (Fragment).

GN SCR.

OS Ethmostigmus rubripes.

OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;

OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Ethmostigmus.

OX NCBI_TaxID=62613;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97411147; PubMed=9259556;

RA Grenier J.K., Garber T.L., Warren R., Whittington P.M., Carroll S.;

RT "Evolution of the entire arthropod Hox gene set predates the origin and radiation of the onychophoran/arthropod clade.";

RL Curr. Biol. 7:547-553 (1997).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AF010178; AAB91392.1; -.

DR HSSP; P02833; 1AHD.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003700; P:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS00071; HOMEBOX 2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON_TER 1

FT SEQUENCE 77 AA; 9377 MW; DF13885AD5B5EB71 CRC64;

Query Match 83.9%; Score 99; DB 5; Length 77;

Best Local Similarity 90.0%; Pred. No. 9.4e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKMKWK 20

Db 33 CLSERQIKIWFQNRKMKWK 52

RESULT 12

Q9PVR7 PRELIMINARY; PRT; 81 AA.

AC Q9PVR7

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE HOXB5B (Fragment).

GN HOXB5B

OS Oryzias latipes (Medaka fish) (Japanese ricefish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

OX NCBI_TaxID=8090;

RN [1]

RP SEQUENCE FROM N.A.

RA Kondo S., Naruse K., Shima A.;

RT "Hox genes of the medakafish Oryzias latipes.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AB026592; BAA86245.1; -.

DR HSSP; P02833; 1SAN.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003700; P:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.

DR PRINTS; PR00024; HOMEBOX.

DR ProDom; PD000010; Homeobox; 1.

DR SMART; SM00010; Homeobox; 1.

DR SMART; SM00389; HOX; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1.

DR PROSITE; PS00071; HOMEBOX 2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

FT NON_TER 1

FT SEQUENCE 81 AA; 9816 MW; 493AFABC783D9BEB CRC64;

Query Match 83.9%; Score 99; DB 13; Length 81;

Best Local Similarity 90.0%; Pred. No. 9.8e-08;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKMKWK 20

Db 44 CLSERQIKIWFQNRKMKWK 63

RESULT 13

Q967W7 PRELIMINARY; PRT; 88 AA.

AC Q967W7

DT 01-DEC-2001 (TRENBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Sex combs reduced (Fragment).
 OS *Polsomia candida*.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Collembola; Arthropoleona;
 OC Entomobryodea; Isotomidae; Proisotominae; Polsomia.
 OX NCBI_TaxID=158441;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21272202; PubMed=11378385;
 RA Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M.E.;
 RT "Hox genes and phylogeny of the arthropods.";
 RL Curr. Biol. 11:759-763 (2001).
 DR EMBL; AF361330; AAK51944.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 FT NON_TER 1
 FT NON_TER 88
 FT NON_TER 88
 SQ SEQUENCE 88 AA; 10889 MW; 2D80B48008E6EC70 CRC64;

 Query Match 83.9%; Score 99; DB 5; Length 88;
 Best Local Similarity 90.0%; Pred. No. 1.1e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 CSSCQIKIWFQNRKMKWK 20
 Db 46 CFSERQIKIWFQNRKMKWK 65

 RESULT 14
 Q967V4 ID Q967V4 PRELIMINARY; PRT; 89 AA.
 AC Q967V4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Sex combs reduced 2 (Fragment).
 GN SCR2.
 OS Lithobius forficatus.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
 OX NCBI_TaxID=7552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21272202; PubMed=11378385;
 RA Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M.E.;
 RT "Hox genes and phylogeny of the arthropods.";
 RL Curr. Biol. 11:759-763 (2001).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF362089; AAK51944.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 89
 SQ SEQUENCE 89 AA; 10703 MW; 240F5DAA5307AC2C CRC64;

 Query Match 83.9%; Score 99; DB 5; Length 89;
 Best Local Similarity 90.0%; Pred. No. 1.1e-07;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 CSSCQIKIWFQNRKMKWK 20
 Db 45 CLSERQIKIWFQNRKMKWK 64

 RESULT 15
 O57377 ID O57377 PRELIMINARY; PRT; 92 AA.
 AC O57377;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hoxv6 protein (Fragment).
 GN HoxC6B OR HOXV6.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Prince V.E., Joly L., Ekker M., Ho R.K.;
 RT "zebrafish hox genes: genomic organization and modified colinear
 expression patterns in the trunk.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; Y14537; CAA74872.1; -.
 DR HSSP; P02833; 9ANT.
 DR ZFIN; ZDB-GENE-000822-1; hoxc6b.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 FT NON_TER 1
 FT NON_TER 92
 SQ SEQUENCE 92 AA; 11150 MW; FDSFF520D2D2641D CRC64;

 Query Match 83.9%; Score 99; DB 13; Length 92;
 Best Local Similarity 90.0%; Pred. No. 1.1e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 CSSCQIKIWFQNRKMKWK 20
 Db 38 CLSERQIKIWFQNRKMKWK 57

 RESULT 16
 Q967V6 ID Q967V6 PRELIMINARY; PRT; 94 AA.
 AC Q967V6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Deformed (Fragment).
 GN DFD.
 OS Lithobius forficatus.
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
 OC Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
 OX NCBI_TaxID=7552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21272202; PubMed=11378385;
 RA Cook C.E., Smith M.L., Telford M.J., Bastianello A., Akam M.E.;
 RT "Hox genes and phylogeny of the arthropods.";
 RL Curr. Biol. 11:759-763 (2001).

RESULT 18
Q94575
ID Q94575 PRELIMINARY;
PRT: 109 AA.

[1] —
SEQUENCE FROM N.A.
STRAIN=North arm of Great Salt Lake;
Averof M., Akam M.;
"HOM/HOX genes in a crustacean: implication for the origin of insect
and crustacean body plans."
Curr. Biol. 3:73-78 (1993).
-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
EMBL: X70080; CAA49685.1; -.
HSSP: P02833; 1SAN.
GO: GO:0005634; C:nucleus; IEA.
GO: GO:0003700; F:transcription factor activity; IEA.
GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
InterPro: IPR001356; Homeobox.
Pfam: PF00046; homeobox; 1.
PRINTS: PR00024; HOMEBOX.

```

DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 109 AA; 13596 MW; 4721740C9D62993B CRC64;

Query Match      83.9%; Score 99; DB 5; Length 109;
Best Local Similarity 90.0%; Pred. No. 1.3e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRMRKWK 20
DB 46 CLSERQIKIWFQNRMRKWK 65

RESULT 20
ID O44258 PRELIMINARY; PRT; 115 AA.
AC O44258;
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Deformed (Fragment).
DN DFD.
OS Ethmostigmus rubripes.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Ethmostigmus.
OX NCBI_TaxID=62613;
RN 1|-----|
RP SEQUENCE FROM N.A.
RX MEDLINE=97411147; PubMed=9259556;
RA Grenier J.K., Garber T.L., Warren R., Whitington P.M., Carroll S.;
RT "Evolution of the entire arthropod Hox gene set predated the origin
RT and radiation of the onychophoran/arthropod clade.";
RL Curr. Biol. 7:547-553 (1997)
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF010176; AAB91390.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 115 AA; 13633 MW; D0AEEC1369ACF517 CRC64;

Query Match      83.9%; Score 99; DB 5; Length 115;
Best Local Similarity 90.0%; Pred. No. 1.4e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRMRKWK 20
DB 19 CLSERQIKIWFQNRMRKWK 38

RESULT 21
ID Q801B5 PRELIMINARY; PRT; 115 AA.
AC Q801B5;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE HoxC4 (Fragment).
OS Latimeria menadoensis (Indonesian coelacanth).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Coelacanthiformes; Coelacanthidae; Latimeria.
RN NCBI_TaxID=106881;
RP SEQUENCE FROM N.A.
RX MEDLINE=22457206; PubMed=12547909;
RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RT "Hox gene clusters in the Indonesian coelacanth, Latimeria
RT menadoensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1084-1088 (2003).
RN 2|-----|
RP SEQUENCE FROM N.A.
RA Koh E.G.L., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
RA Venkatesh B.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY183742; AAO43035.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 13462 MW; 956057D9822E9B52 CRC64;

Query Match      83.9%; Score 99; DB 13; Length 115;
Best Local Similarity 90.0%; Pred. No. 1.4e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCQIKIWFQNRMRKWK 20
DB 47 CLSERQIKIWFQNRMRKWK 66

RESULT 22
ID Q05009 PRELIMINARY; PRT; 146 AA.
AC Q05009;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Dfd protein (Fragment).
DN DFD.
OS Artemia sanfranciscana (Brine shrimp) (Artemia franciscana).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OX NCBI_TaxID=6661;
RN 1|-----|
RP SEQUENCE FROM N.A.
RC STRAIN=North arm of Great Salt Lake;
RA Averof M., Akam M.;
RT "HOM/HOX genes in a crustacean: implication for the origin of insect
RT and crustacean body plans.";
RL Curr. Biol. 3:73-78 (1993).
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; X70078; CAA49683.1; -.
DR HSSP; P02833; 9ANT.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 1

```

SQ SEQUENCE 146 AA; 16972 MW; 03D47AD09575C129 CRC64;

Query Match 83.9%; Score 99; DB 5; Length 146;
 Best Local Similarity 90.0%; Pred. No. 1.7e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKWK 20
 DB 49 CLSERQIKIWFQNRKWK 68

RESULT 23

Q61679 PRELIMINARY; PRT; 150 AA.

AC Q61679;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Homeobox protein (Hox-2.6) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Liver;
 RX MEDLINE=89138440; PubMed=2906328;
 RA Do M.S., Lonai P.;
 RT "Gene organization of murine homeobox-containing gene clusters.";
 RL Genomics 3:195-200(1988).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; M18342; AAY75162.1; -.
 DR EIR; A27348; A27348.
 DR HSP; P02833; 9ANT.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; Homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 KW NON_TER 1
 FT NON_TER 150 150
 SQ SEQUENCE 150 AA; 16843 MW; 156B423971665D78 CRC64;

Query Match 83.9%; Score 99; DB 11; Length 150;
 Best Local Similarity 90.0%; Pred. No. 1.7e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKWK 20
 DB 73 CLSERQIKIWFQNRKWK 92

RESULT 24

Q801C0 PRELIMINARY; PRT; 225 AA.

AC Q801C0;

DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE HoxB5.
 OS Latimeria menadoensis (Indonesian coelacanth).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Coelacanthiformes; Coelacanthidae; Latimeria.
 OX NCBI_TaxID=106881;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22457206; PubMed=12547909;

RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
 RA Venkatesh B.;
 RT "Hox gene clusters in the Indonesian coelacanth, Latimeria
 RT menadoensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1084-1088(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Koh E.G., Lam K., Christoffels A., Erdmann M.V., Brenner S.,
 RA Venkatesh B.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY183737; AAC43030.1; -.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 SQ SEQUENCE 225 AA; 25098 MW; 6F8A24B2149D5661 CRC64;

Query Match 83.9%; Score 99; DB 13; Length 225;
 Best Local Similarity 90.0%; Pred. No. 2.5e-07;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCRQIKIWFQNRKWK 20
 DB 188 CLSERQIKIWFQNRKWK 207

RESULT 25

Q9YGT6 PRELIMINARY; PRT; 227 AA.

AC Q9YGT6;

DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Homeobox protein.
 GN HOXA5A.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99051425; PubMed=9831563;
 RA Amores A., Force A., Yan Y.-L., Wang Y.-L., Fritz A., Prince V.,
 RA Ho R., Amemiya C., Langeland J., Westerfield M., Ekker M.,
 RA Postlethwait J.;
 RT "Zebrafish hox clusters and vertebrate genome evolution.";
 RL Science 282:1711-1714(1998).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AF071247; AAD15940.1; -.
 DR HSP; P02833; 1SAN.
 DR TRANSFAC; T03620; -.
 DR ZFIN; ZDB-GENE-000823-9; hoxa5a.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00025; ANTENNAPEDIA.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.

DR PROSITE; PS00027; HOMEBOX 1; 1;
DR PROSITE; PS50071; HOMEBOX_2; 1;
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 227 AA; 25310 MW; 608CAFBA72B8D1F CRC64;
Query Match 83.9%; Score 99; DB 13; Length 227;
Best Local Similarity 90.0%; Pred. No. 2.6e-07;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CSSCRQIKIWFQNRMKKK 20
Db 190 CLSERQIKIWFQNRMKKK 209

Search completed: May 24, 2004, 17:24:40
Job time : 34.5946 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:22:27 ; Search time 14.3243 Seconds
(without alignments)
72.082 Million cell updates/sec

Title: US-09-977-349-3

Perfect score: 118

Sequence: 1 CSSCRQIKINFQNRMMKKW 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listings first 100 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	101	85.6	17	4	US-09-346-847-17
2	101	85.6	17	4	US-09-346-847-27
3	101	85.6	20	4	US-09-346-847-30
4	99	83.9	61	3	US-08-751-344B-9
5	98	83.1	20	4	US-09-658-517C-8
6	98	83.1	61	3	US-08-751-344B-7
7	96	81.4	61	2	US-08-202-044-3
8	96	81.4	61	3	US-08-751-344B-3
9	96	81.4	61	3	US-08-751-344B-6
10	93	78.8	24	4	US-09-428-082B-332
11	92	78.0	16	2	US-08-928-958-7
12	92	78.0	16	2	US-08-810-540-3
13	92	78.0	16	2	US-09-072-429-7
14	92	78.0	16	3	US-08-964-302A-6
15	92	78.0	16	3	US-09-116-294-4
16	92	78.0	16	3	US-08-964-614A-4
17	92	78.0	16	3	US-08-849-486-1
18	92	78.0	16	3	US-08-849-486-4
19	92	78.0	16	3	US-09-208-966-54
20	92	78.0	16	3	US-09-308-935-8
21	92	78.0	16	3	US-09-441-416A-6
22	92	78.0	16	4	US-09-296-089-33
23	92	78.0	16	4	US-09-419-826-35
24	92	78.0	16	4	US-09-302-305C-10
25	92	78.0	16	4	US-09-346-847-1
26	92	78.0	16	4	US-09-346-847-1
27	92	78.0	16	4	US-09-346-847-25

28	92	78.0	16	4	US-09-057-363C-47	Sequence 47, Appli
29	92	78.0	16	4	US-09-043-560B-3	Sequence 3, Appli
30	92	78.0	16	4	US-09-648-400A-29	Sequence 29, Appli
31	92	78.0	16	4	US-09-227-652B-4	Sequence 4, Appli
32	92	78.0	16	4	US-09-780-070-38	Sequence 38, Appli
33	92	78.0	16	4	US-08-610-220B-9	Sequence 9, Appli
34	92	78.0	16	4	US-09-775-052A-54	Sequence 54, Appli
35	92	78.0	16	4	US-09-155-165-22	Sequence 22, Appli
36	92	78.0	16	4	US-09-792-480-29	Sequence 29, Appli
37	92	78.0	16	4	US-09-551-976-33	Sequence 33, Appli
38	92	78.0	16	4	US-09-265-107-47	Sequence 47, Appli
39	92	78.0	16	4	US-09-346-847-20	Sequence 20, Appli
40	92	78.0	17	4	US-09-346-847-22	Sequence 22, Appli
41	92	78.0	17	4	US-09-648-400A-30	Sequence 30, Appli
42	92	78.0	18	3	US-08-838-545-20	Sequence 20, Appli
43	92	78.0	18	3	US-09-349-532-20	Sequence 20, Appli
44	92	78.0	19	4	US-09-346-847-23	Sequence 23, Appli
45	92	78.0	19	4	US-09-658-517C-7	Sequence 7, Appli
46	92	78.0	20	4	US-09-346-847-16	Sequence 16, Appli
47	92	78.0	20	4	US-09-346-847-18	Sequence 18, Appli
48	92	78.0	21	4	US-08-610-220B-11	Sequence 11, Appli
49	92	78.0	22	4	US-09-346-847-28	Sequence 28, Appli
50	92	78.0	22	4	US-09-057-363C-50	Sequence 50, Appli
51	92	78.0	22	4	US-08-610-220B-10	Sequence 10, Appli
52	92	78.0	22	4	US-09-155-165-5	Sequence 5, Appli
53	92	78.0	22	4	US-09-155-165-11	Sequence 11, Appli
54	92	78.0	24	4	US-09-265-107-50	Sequence 50, Appli
55	92	78.0	24	4	US-09-419-826-34	Sequence 34, Appli
56	92	78.0	27	3	US-09-051-934-51	Sequence 51, Appli
57	92	78.0	27	3	US-09-051-934-52	Sequence 52, Appli
58	92	78.0	27	4	US-09-040-725A-2	Sequence 2, Appli
59	92	78.0	34	4	US-09-347-504-79	Sequence 79, Appli
60	92	78.0	34	4	US-10-161-499-79	Sequence 79, Appli
61	92	78.0	36	4	US-09-428-082B-331	Sequence 331, App
62	87	73.7	15	2	US-08-810-540-4	Sequence 4, Appli
63	87	73.7	16	4	US-09-792-480-30	Sequence 30, Appli
64	87	73.7	42	3	US-08-751-344B-4	Sequence 4, Appli
65	87	73.7	283	1	US-08-583-672-2	Sequence 2, Appli
66	87	73.7	283	2	US-08-202-044-2	Sequence 2, Appli
67	87	73.7	283	3	US-08-751-344B-2	Sequence 2, Appli
68	87	73.7	284	2	US-08-320-148B-2	Sequence 2, Appli
69	87	73.7	284	3	US-08-589-028-6	Sequence 6, Appli
70	87	73.7	284	3	US-08-784-582-6	Sequence 6, Appli
71	87	73.7	284	3	US-08-785-271-6	Sequence 6, Appli
72	87	73.7	284	3	US-09-031-898-2	Sequence 2, Appli
73	87	73.7	302	2	US-08-203-532F-4	Sequence 4, Appli
74	87	73.7	302	3	US-08-950-860-16	Sequence 16, Appli
75	87	73.7	302	3	US-09-078-485-4	Sequence 4, Appli
76	87	73.7	302	5	PCT-US95-01882A-4	Sequence 4, Appli
77	87	73.7	303	2	US-08-203-532F-2	Sequence 2, Appli
78	87	73.7	303	3	US-09-078-465-2	Sequence 2, Appli
79	87	73.7	303	5	PCT-US95-01882A-2	Sequence 2, Appli
80	86	72.9	16	3	US-08-849-486-5	Sequence 5, Appli
81	86	72.9	16	4	US-09-466-772-7	Sequence 7, Appli
82	86	72.9	20	4	US-09-466-772-3	Sequence 3, Appli
83	86	72.9	21	4	US-09-466-772-1	Sequence 1, Appli
84	86	72.9	22	4	US-09-466-772-2	Sequence 2, Appli
85	86	72.9	23	4	US-09-466-772-4	Sequence 4, Appli
86	83	70.3	16	3	US-08-849-486-8	Sequence 8, Appli
87	83	70.3	300	4	US-09-162-524-1	Sequence 1, Appli
88	83	70.3	349	4	US-09-162-524-3	Sequence 3, Appli
89	82	69.5	16	4	US-09-404-929-7	Sequence 7, Appli
90	80	67.9	16	3	US-08-849-486-7	Sequence 7, Appli
91	79	66.8	61	2	US-08-757-316C-30	Sequence 30, Appli
92	78	66.1	16	3	US-09-208-966-1	Sequence 1, Appli
93	78	66.1	16	4	US-09-775-052A-1	Sequence 1, Appli
94	78	66.1	342	4	US-09-165-827C-2	Sequence 2, Appli
95	77	65.3	60	3	US-08-751-344B-5	Sequence 5, Appli
96	76	64.4	367	4	US-09-009-816-2	Sequence 2, Appli
97	75	63.6	16	4	US-09-296-089-34	Sequence 34, Appli
98	75	63.6	16	4	US-09-336-093-3	Sequence 3, Appli
99	75	63.6	16	4	US-09-057-363C-48	Sequence 48, Appli
100	75	63.6	16	4	US-09-557-465D-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-346-847-17
; Sequence 17, Application US/09346847
; Patent No. 6472507
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/09/346,847
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-346-847-17

Query Match 85.6%; Score 101; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CQIKIWFQNRMKWKX 20
|||
Db 1 CQIKIWFQNRMKWKX 17

RESULT 2
US-09-346-847-27
; Sequence 27, Application US/09346847
; Patent No. 6472507
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/09/346,847
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (17)
; OTHER INFORMATION: AMIDATION
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-346-847-27

Query Match 85.6%; Score 101; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CQIKIWFQNRMKWKX 20
|||
Db 1 CQIKIWFQNRMKWKX 17

RESULT 3
US-09-346-847-30
; Sequence 30, Application US/09346847
; Patent No. 6472507
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/09/346,847
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-346-847-30

Query Match 85.6%; Score 101; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CQIKIWFQNRMKWKX 20
|||
Db 1 CQIKIWFQNRMKWKX 17

RESULT 4
US-08-751-344B-9
; Sequence 9, Application US/08751344B
; Patent No. 6210960
; GENERAL INFORMATION:
; APPLICANT: Habener M.D., Joel P.
; APPLICANT: Miller Ph.D., Christopher P.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,344B
; FILING DATE: 19-NO. 6210960-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/202,044
; FILING DATE: 23-Feb-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid

```

/ COMPUTER MODEL: PC-AT
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WordPerfect 6.1
/ CURRENT APPLICATION DATA:
/

```

HYPOTHETICAL: NO
ANTI-SENSE: NO

```
/ FRAGMENT TYPE: internal
US-08-202-044-3

Query Match      81.4%; Score 96; DB 2; Length 61;
Best Local Similarity 85.0%; Pred. No. 8.3e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCROIKWFOQNRMMKWK 20
Db 40 CLTERQIKWFOQNRMMKWK 59

RESULT 8
US-08-751-344B-3
; Sequence 3, Application US/08751344B
; Patent No. 6210960
; GENERAL INFORMATION:
; APPLICANT: Habener M.D., Joel F.
; TITLE OF INVENTION: NOVEL TRANSCRIPTION FACTOR AND USES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,344B
; FILING DATE: 19-No. 6210960-1996
; PRIOR APPLICATION DATA: 08/202,044
; FILING DATE: 23-Feb-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,137-A (11274/02148)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-751-344B-6

Query Match      81.4%; Score 96; DB 3; Length 61;
Best Local Similarity 85.0%; Pred. No. 8.3e-07;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CSSCROIKWFOQNRMMKWK 20
Db 40 CLTERQIKWFOQNRMMKWK 59

RESULT 10
US-09-428-082B-332
; Sequence 332, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
US-09-428-082B-332
```



```

Query Match      78.8%; Score 93; DB 4; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.7e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  SCROIKWIFQNRMRKWK 20
      |  |||||
Db       7  SKROIKWIFQNRMRKWK 24

RESULT 11
US-08-928-958-7
; Sequence 7, Application US/08928958
; Patent No. 5877282
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CLEAVELAND, JEFFREY S.
; APPLICANT: BLAKE, JAMES
; APPLICANT: HAFAR, OMAR K.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,958
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026978
; FILING DATE: 20-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-928-958-7

Query Match      78.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  RQIKWIFQNRMRKWK 20
      |  |||||
Db       1  RQIKWIFQNRMRKWK 16

RESULT 12
US-08-810-540-3
; Sequence 3, Application US/08810540
; Patent No. 5929042
; GENERAL INFORMATION:
; APPLICANT: TROY, CAROL M.

```

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/51247
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-810-540-6

Query Match 78.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQIKWFWQRMRMVKK 20
Db 1 RQIKWFWQRMRMVKK 16

RESULT 14
US-09-072-429-7
Sequence 7, Application US/09072429
Patent No. 5962415
GENERAL INFORMATION:
APPLICANT: Nadler, Steven G.
TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,429
FILING DATE: 04-MAY-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Klein, Christopher A.
REGISTRATION NUMBER: 34,363
REFERENCE/DOCKET NUMBER: ON0141b
TELEPHONE: (609) 252-3714
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-072-429-7

Query Match 78.0%; Score 92; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQIKWFWQRMRMVKK 20
Db 1 RQIKWFWQRMRMVKK 16

RESULT 15
US-08-964-302A-6
Sequence 6, Application US/08964302A
Patent No. 6015787
GENERAL INFORMATION:
APPLICANT: Potter, David A.
APPLICANT: Skolnik, Paul R.
TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF CALPAIN
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,302A
FILING DATE: 04-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 00398/126001
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-964-302A-6

Query Match 78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQIKWFWQRMRMVKK 20
Db 1 RQIKWFWQRMRMVKK 16

RESULT 16
US-09-116-294-4
Sequence 4, Application US/09116294
Patent No. 6025140
GENERAL INFORMATION:
APPLICANT: Langel, Ulo
APPLICANT: Bartfai, Tamas
APPLICANT: Pooga, Margus
APPLICANT: Valkna, Andres
APPLICANT: Saar, Kulliki
APPLICANT: Hallbrink, Mattias
TITLE OF INVENTION: Conjugated Constructs of Peptides and
TITLE OF INVENTION: Nucleic Acid Analogs, and Their Transport Across Membranes
FILE REFERENCE: 4394
CURRENT APPLICATION NUMBER: US/09/116,294
CURRENT FILING DATE: 1998-07-16
EARLIER APPLICATION NUMBER: 60/052,678
EARLIER FILING DATE: 1997-07-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 16
TYPE: PRT

```
; ORGANISM: drosophila
US-09-116-294-4

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWQNRRMKWKK 20
DB      1 RQIKWQNRRMKWKK 16
|||||

RESULT 17
US-08-964-614A-4
; Sequence 4, Application US/08964614A
; Patent No. 6057104
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN
; TITLE OF INVENTION: Rad51 PROTEIN AND DISRUPTION OF PROTEINS THAT ASSOCIATE
; TITLE OF INVENTION: WITH MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,614A
; FILING DATE: 05-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,280
; FILING DATE: 05-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8535-0019-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-964-614A-4

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWQNRRMKWKK 20
DB      1 RQIKWQNRRMKWKK 16
|||||

RESULT 18
US-08-849-486-1
; Sequence 1, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..16
; OTHER INFORMATION: /product= "amino acids of the D series"
US-08-849-486-4

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWQNRRMKWKK 20
DB      1 RQIKWQNRRMKWKK 16
|||||

RESULT 19
US-08-849-486-4
; Sequence 4, Application US/08849486
; Patent No. 6080724
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDES WHICH CAN BE USED AS VECTORS
; TITLE OF INVENTION: FOR THE INTRACELLULAR ADDRESSING OF ACTIVE MOLECULES
; NUMBER OF SEQUENCES: 10
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,486
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 11714
; FILING DATE: 05-OCT-1995
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..16
; OTHER INFORMATION: /product= "amino acids of the D series"
US-08-849-486-4

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWQNRRMKWKK 20
DB      1 RQIKWQNRRMKWKK 16
|||||
```

```
Db      1 RQIKWIFQNRMRMKWK 16

RESULT 20
US-09-208-966-54
; Sequence 54, Application US/09208966
; Patent No. 6221355
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 54
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-208-966-54

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRMRMKWK 20
        |||||
Db      1 RQIKWIFQNRMRMKWK 16

RESULT 21
US-09-308-935-8
; Sequence 8, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-308-935-8

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRMRMKWK 20
        |||||
Db      1 RQIKWIFQNRMRMKWK 16

RESULT 22
US-09-441-416A-6
; Sequence 6, Application US/09441416A
; Patent No. 6294518
; GENERAL INFORMATION:
; APPLICANT: Potter, David A.
; APPLICANT: Skolnik, Paul R.
; TITLE OF INVENTION: CELL-PERMEABLE PROTEIN INHIBITORS OF
; FILE REFERENCE: 00398-140001
; CURRENT APPLICATION NUMBER: US/09/441,416A
; CURRENT FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: US 08/964,302
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-441-416A-6

Query Match      78.0%; Score 92; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRMRMKWK 20
        |||||
Db      1 RQIKWIFQNRMRMKWK 16

RESULT 23
US-09-296-089-33
; Sequence 33, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; FILE REFERENCE: 10086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-296-089-33

Query Match      78.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRMRMKWK 20
        |||||
Db      1 RQIKWIFQNRMRMKWK 16

RESULT 24
US-09-419-826-35
; Sequence 35, Application US/09419826
; Patent No. 6306832
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PEPTIDE ANTISTROGEN COMPOSITIONS AND METHODS
; FILE REFERENCE: 100.0%; Pred. No. 8.1e-07;
; NUMBER OF SEQUENCES: 39
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/09/419,826
; FILING DATE: 14-OCT-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/07711
```

```
; FILING DATE: 14-APR-1998
; APPLICATION NUMBER: US 60/043,545
; FILING DATE: 14-APR-1997
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-09-419-826-35

Query Match          78.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRRMKWK 20
Db      1 RQIKWIFQNRRMKWK 16

RESULT 25
US-09-302-305C-10
; Sequence 10, Application US/09302305C
; Patent No. 6350572
; GENERAL INFORMATION:
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Interaction Between Cyclin D1 and Steroid Receptor
; FILE REFERENCE: 4238/80713
; CURRENT APPLICATION NUMBER: US/09/302,305C
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: PCT/GB99/00440
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(16)
; OTHER INFORMATION: Translocation peptide derived from antennapedia
; OTHER INFORMATION: homeodomain protein
;
US-09-302-305C-10

Query Match          78.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 RQIKWIFQNRRMKWK 20
Db      1 RQIKWIFQNRRMKWK 16

Search completed: May 24, 2004, 17:29:50
Job time : 15.3243 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:19:01 ; Search time 37.5676 Seconds

(without alignments)
148.491 Million cell updates/sec

Title: US-09-977-349-3

Perfect score: 118

Sequence: 1 CGSQRQIKWQFQNRKMKKK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09H_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	101	85.6	17	12	US-10-421-548-8
2	101	85.6	17	13	US-10-007-761-8
3	101	85.6	17	14	US-10-210-660-17
4	101	85.6	17	14	US-10-210-660-27
5	101	85.6	17	15	US-10-372-003A-29
6	101	85.6	17	15	US-10-428-280-15
7	101	85.6	17	15	US-10-421-503-66
8	101	85.6	20	14	US-09-854-204-63
9	101	85.6	20	14	US-10-116-275-190
10	99	83.5	269	15	US-10-176-419A-4
11	98.5	83.5	29	15	US-09-949-474-8
12	98	83.1	20	9	US-09-949-474-8
13	96	81.4	64	14	US-10-118-079-44
14	96	81.4	217	12	US-10-097-105-1561
15	96	81.4	217	14	US-10-097-340-129

16	96	81.4	233	15	US-10-420-940-4	Sequence 4, Appli
17	96	81.4	235	14	US-10-118-079-4	Sequence 4, Appli
18	93	78.8	24	12	US-10-609-217-332	Sequence 332, App
19	93	78.8	24	12	US-10-632-388-332	Sequence 332, App
20	93	78.8	24	12	US-10-651-723-332	Sequence 332, App
21	93	78.8	24	12	US-10-645-761-332	Sequence 332, App
22	93	78.8	24	16	US-10-666-696-332	Sequence 332, App
23	93	78.8	24	16	US-10-653-048-332	Sequence 332, App
24	92	78.0	16	8	US-08-610-220A-9	Sequence 9, Appli
25	92	78.0	16	9	US-09-214-371-43	Sequence 43, Appli
26	92	78.0	16	9	US-09-780-070-38	Sequence 38, Appli
27	92	78.0	16	9	US-09-150-623-9	Sequence 9, Appli
28	92	78.0	16	9	US-09-731-023A-10	Sequence 10, Appli
29	92	78.0	16	9	US-09-854-204-1	Sequence 1, Appli
30	92	78.0	16	9	US-09-900-147-8	Sequence 8, Appli
31	92	78.0	16	9	US-09-792-480-29	Sequence 29, Appli
32	92	78.0	16	9	US-09-785-802A-2	Sequence 2, Appli
33	92	78.0	16	9	US-09-785-802A-5	Sequence 5, Appli
34	92	78.0	16	9	US-09-902-432-32	Sequence 32, Appli
35	92	78.0	16	9	US-09-953-031A-10	Sequence 10, Appli
36	92	78.0	16	9	US-09-981-286A-3	Sequence 3, Appli
37	92	78.0	16	10	US-09-962-967A-6	Sequence 6, Appli
38	92	78.0	16	10	US-09-912-414-6	Sequence 6, Appli
39	92	78.0	16	10	US-09-775-052-54	Sequence 54, Appli
40	92	78.0	16	10	US-09-295-189-4	Sequence 4, Appli
41	92	78.0	16	10	US-09-965-876A-1	Sequence 1, Appli
42	92	78.0	16	11	US-09-933-780C-2	Sequence 2, Appli
43	92	78.0	16	12	US-10-286-964-3	Sequence 3, Appli
44	92	78.0	16	12	US-10-432-291-6	Sequence 6, Appli
45	92	78.0	16	12	US-10-357-826A-2	Sequence 2, Appli
46	92	78.0	16	12	US-10-603-409-12	Sequence 12, Appli
47	92	78.0	16	12	US-09-748-083-3	Sequence 3, Appli
48	92	78.0	16	12	US-09-779-791A-3	Sequence 3, Appli
49	92	78.0	16	12	US-10-083-960-29	Sequence 29, Appli
50	92	78.0	16	12	US-10-144-549-1	Sequence 1, Appli
51	92	78.0	16	12	US-10-136-187-21	Sequence 21, Appli
52	92	78.0	16	12	US-10-261-161-1	Sequence 1, Appli
53	92	78.0	16	12	US-10-261-161-26	Sequence 26, Appli
54	92	78.0	16	13	US-10-024-935-12	Sequence 12, Appli
55	92	78.0	16	13	US-10-007-363-3	Sequence 3, Appli
56	92	78.0	16	14	US-10-071-512A-2	Sequence 2, Appli
57	92	78.0	16	14	US-10-239-804-3	Sequence 3, Appli
58	92	78.0	16	14	US-10-077-555-3	Sequence 3, Appli
59	92	78.0	16	14	US-10-209-421-29	Sequence 29, Appli
60	92	78.0	16	14	US-10-229-915-2	Sequence 2, Appli
61	92	78.0	16	14	US-10-185-084-3	Sequence 3, Appli
62	92	78.0	16	14	US-10-252-012-5	Sequence 5, Appli
63	92	78.0	16	14	US-10-075-869-19	Sequence 19, Appli
64	92	78.0	16	14	US-10-013-815-19	Sequence 19, Appli
65	92	78.0	16	14	US-10-136-738-10	Sequence 10, Appli
66	92	78.0	16	14	US-10-210-660-1	Sequence 1, Appli
67	92	78.0	16	14	US-10-210-660-25	Sequence 25, Appli
68	92	78.0	16	14	US-10-156-570A-21	Sequence 21, Appli
69	92	78.0	16	14	US-10-201-394A-14	Sequence 14, Appli
70	92	78.0	16	14	US-10-017-672-11	Sequence 11, Appli
71	92	78.0	16	14	US-10-201-389A-14	Sequence 14, Appli
72	92	78.0	16	14	US-10-161-051-1	Sequence 1, Appli
73	92	78.0	16	14	US-10-358-365-10	Sequence 10, Appli
74	92	78.0	16	14	US-10-061-607A-2	Sequence 2, Appli
75	92	78.0	16	14	US-10-405-339-44	Sequence 44, Appli
76	92	78.0	16	15	US-10-366-493-19	Sequence 19, Appli
77	92	78.0	16	15	US-10-444-662-2	Sequence 2, Appli
78	92	78.0	16	15	US-10-185-593-2	Sequence 2, Appli
79	92	78.0	16	15	US-10-413-160-38	Sequence 38, Appli
80	92	78.0	16	15	US-10-462-138-10	Sequence 10, Appli
81	92	78.0	16	15	US-10-369-226-47	Sequence 47, Appli
82	92	78.0	16	15	US-10-353-678-2	Sequence 2, Appli
83	92	78.0	16	15	US-10-357-529-8	Sequence 8, Appli
84	92	78.0	17	9	US-09-854-204-19	Sequence 19, Appli
85	92	78.0	17	9	US-09-785-802A-3	Sequence 3, Appli
86	92	78.0	17	11	US-09-933-780C-21	Sequence 21, Appli
87	92	78.0	17	12	US-10-602-303-3	Sequence 3, Appli
88	92	78.0	17	14	US-10-209-421-30	Sequence 30, Appli

Sequence 1, Appli
Sequence 20, Appl
Sequence 22, Appl
Sequence 14, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 23, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 11, Appl
Sequence 11, Appl

89 92 78.0 17 14 US-10-229-915-1
90 92 78.0 17 14 US-10-210-660-20
91 92 78.0 17 14 US-10-210-660-22
92 92 78.0 18 9 US-09-785-802A-14
93 92 78.0 19 9 US-09-949-474-7
94 92 78.0 19 14 US-10-118-079-45
95 92 78.0 19 14 US-10-210-660-23
96 92 78.0 19 15 US-10-407-449-20
97 92 78.0 20 14 US-10-210-660-16
98 92 78.0 20 14 US-10-210-660-18
99 92 78.0 21 8 US-08-610-220A-11
100 92 78.0 21 9 US-09-150-623-11

ALIGNMENTS

RESULT 1
US-10-421-548-8
; Sequence 8, Application US/10421548
; Publication No. US20030223981A1
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Sweitzer, Sarah M.
; APPLICANT: Kendig, Joan J.
; APPLICANT: Yeomans, David C.
; TITLE OF INVENTION: Peptide Inhibitors of Protein Kinase C
; TITLE OF INVENTION: Gamma for Pain Management
; FILE REFERENCE: 58600-8210.US00
; CURRENT APPLICATION NUMBER: US/10/421,548
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/374,530
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Drosophila Antennapedia homeodomain-derived
; OTHER INFORMATION: carrier peptide
US-10-421-548-8

Query Match 85.6%; Score 101; DB 12; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CRQIKWFOQRMRMKWK 20
Db 1 CRQIKWFOQRMRMKWK 17

RESULT 2
US-10-007-761-8
; Sequence 8, Application US/10007761
; Publication No. US20020150984A1
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: Peptides for Activation and Inhibition
; TITLE OF INVENTION: of delta-PRC
; FILE REFERENCE: 58600-8208.US00
; CURRENT APPLICATION NUMBER: US/10/007,761
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/262,060
; PRIOR FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Drosophila Antennapedia homeodomain-derived
; OTHER INFORMATION: carrier peptide
US-10-007-761-8

Query Match 85.6%; Score 101; DB 13; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CRQIKWFOQRMRMKWK 20
Db 1 CRQIKWFOQRMRMKWK 17

RESULT 3
US-10-210-660-17
; Sequence 17, Application US/10210660
; Publication No. US20030119735A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/10/210,660
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US/09/346,847
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-210-660-17

Query Match 85.6%; Score 101; DB 14; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CRQIKWFOQRMRMKWK 20
Db 1 CRQIKWFOQRMRMKWK 17

RESULT 4
US-10-210-660-27
; Sequence 27, Application US/10210660
; Publication No. US20030119735A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/10/210,660
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US/09/346,847
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (17)
; OTHER INFORMATION: AMIDATION

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-210-660-27

Query Match 85.6%; Score 101; DB 14; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRQIKIWFQNRMRMKWK 20
Db 1 CRQIKIWFQNRMRMKWK 17

RESULT 5

US-10-372-003A-29
; Sequence 29, Application US/10372003A
; Publication No. US20030215846A1
; GENERAL INFORMATION:
; APPLICANT: Watt, Paul
; APPLICANT: Thomas, Wayne
; APPLICANT: Hopkins, Richard
; TITLE OF INVENTION: Methods of constructing and screening
; TITLE OF INVENTION: diverse expression libraries
; FILE REFERENCE: FBRT40.001CPI
; CURRENT APPLICATION NUMBER: US/10/372,003A
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/568,229
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/132,711
; PRIOR FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Drosophila
US-10-372-003A-29

Query Match 85.6%; Score 101; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRQIKIWFQNRMRMKWK 20
Db 1 CRQIKIWFQNRMRMKWK 17

RESULT 6

US-10-428-280-15
; Sequence 15, Application US/10428280
; Publication No. US2004000919A1
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Kendig, Joan J.
; APPLICANT: Sweitzer, Sarah M.
; TITLE OF INVENTION: Protein Kinase C Peptides for Use in Withdrawal
; FILE REFERENCE: 58600-8211.US00
; CURRENT APPLICATION NUMBER: US/10/428,280
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US 60/377,331
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Drosophila
; OTHER INFORMATION: Antennapedia homeodomain-derived carrier peptide
US-10-428-280-15

Query Match 85.6%; Score 101; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRQIKIWFQNRMRMKWK 20
Db 1 CRQIKIWFQNRMRMKWK 17

RESULT 7

US-10-421-503-66
; Sequence 66, Application US/10421503
; Publication No. US2004000922A1
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; TITLE OF INVENTION: Peptide inhibitors of protein kinase C
; FILE REFERENCE: 58600-8210.US01
; CURRENT APPLICATION NUMBER: US/10/421,503
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/374,530
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Drosophila Antennapedia homeodomain-derived carrier peptide
US-10-421-503-66

Query Match 85.6%; Score 101; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRQIKIWFQNRMRMKWK 20
Db 1 CRQIKIWFQNRMRMKWK 17

RESULT 8

US-09-854-204-63
; Sequence 63, Application US/09854204
; Patent No. US20020098236A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Peter Martin
; APPLICANT: Zhelev, Nikolai
; TITLE OF INVENTION: Transport Vectors
; FILE REFERENCE: CCI-010
; CURRENT APPLICATION NUMBER: US/09/854,204
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/438,460
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: GB 9825000.4
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9825001.2
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9902525.6
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9902522.3
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: GB 9914578.1
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCT/GB99/03750
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES

; LOCATION: (20)
; OTHER INFORMATION: AMIDATION
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: sequence
US-09-854-204-63

Query Match 85.6%; Score 101; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CROIKIWFQNRMRKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 CROIKIWFQNRMRKWK 17

RESULT 9
US-10-210-660-30
; Sequence 30, Application US/10210660
; Publication No. US20030119735A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, M. Peter
; APPLICANT: Wang, Shudong
; TITLE OF INVENTION: Delivery System
; FILE REFERENCE: CCI-009
; CURRENT APPLICATION NUMBER: US/10/210,660
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US/09/346,847
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: GB 9814527
; PRIOR FILING DATE: 1998-07-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
US-10-210-660-30

Query Match 85.6%; Score 101; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CROIKIWFQNRMRKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 CROIKIWFQNRMRKWK 17

RESULT 10
US-10-116-275-190
; Sequence 190, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 190
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-190

Query Match 83.9%; Score 99; DB 15; Length 269;
Best Local Similarity 90.0%; Pred. No. 1.2e-05;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMRKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 232 CLSERIKIWFQNRMRKWK 251

RESULT 11
US-10-176-419A-4
; Sequence 4, Application US/10176419A
; Publication No. US20040006203A1
; GENERAL INFORMATION:
; APPLICANT: Maier, Martin A.
; APPLICANT: Guzaev, Andrei P.
; APPLICANT: Manoharan, Muthiah
; TITLE OF INVENTION: Method For Solid Phase Synthesis Of PNA Conjugates Using Branched
; TITLE OF INVENTION: Bridging Units Involving Orthogonal Protecting Groups
; FILE REFERENCE: ISIS0507
; CURRENT APPLICATION NUMBER: US/10/176,419A
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide sequence
US-10-176-419A-4

Query Match 83.5%; Score 98.5; DB 15; Length 29;
Best Local Similarity 66.7%; Pred. No. 1.9e-06;
Matches 18; Conservative 2; Mismatches 0; Indels 7; Gaps 1;

QY 1 CSSC-----RQIKIWFQNRMRKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 3 CTTCCCKKKRQIKIWFQNRMRKWK 29

RESULT 12
US-09-949-474-8
; Sequence 8, Application US/09949474
; Patent No. US20020156235A1
; GENERAL INFORMATION:
; APPLICANT: Guzaev, Andrei P.
; APPLICANT: Manoharan, Muthiah
; TITLE OF INVENTION: Process for Preparing Peptide Derivatized Oligomeric Compounds
; FILE REFERENCE: ISIS4850
; CURRENT APPLICATION NUMBER: US/09/949,474
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/658,517
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020156235A1el Sequence
; NAME/KEY: misc feature
; LOCATION: (2)..(2)
; OTHER INFORMATION: Xaa is aminobutyric acid
US-09-949-474-8

Query Match 83.1%; Score 98; DB 9; Length 20;
Best Local Similarity 85.0%; Pred. No. 1.6e-06;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMRKWK 20

RESULT 15

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-420-940-4

Query Match      81.4%; Score 96; DB 15; Length 233;
Best Local Similarity 85.0%; Pred. No. 2.7e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMKWK 20
Db 193 CLTERQIKIWFQNRMKWK 212

RESULT 17
US-10-118-079-4
; Sequence 4, Application US/10118079
; Publication No. US20030103957A1
; GENERAL INFORMATION:
; APPLICANT: MCKERRACHER, LISA
; TITLE OF INVENTION: FUSION PROTEINS
; FILE REFERENCE: 06746-004-US-03
; CURRENT APPLICATION NUMBER: US/10/118,079
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: CA 2,367,636
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: CA 2,362,004
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: CA 2,342,970
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of C3APL: includes ADP-ribosyl transferase C3 (Clostrid
; OTHER INFORMATION: ium botulinum) and Antennapedia sequence.
US-10-118-079-4

Query Match      81.4%; Score 96; DB 14; Length 295;
Best Local Similarity 85.0%; Pred. No. 3.3e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CSSCROIKIWFQNRMKWK 20
Db 274 CLTERQIKIWFQNRMKWK 293

RESULT 18
US-10-609-217-332
; Sequence 332, Application US/10609217
; Publication No. US20040044188A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/609,217
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
```

```
; OTHER INFORMATION: P16-MIMETIC
US-10-609-217-332

Query Match      78.8%; Score 93; DB 12; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SCROIKIWFQNRMKWK 20
Db 7 SKROIKIWFQNRMKWK 24

RESULT 19
US-10-632-388-332
; Sequence 332, Application US/10632388
; Publication No. US20040053845A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/632,388
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
US-10-632-388-332

Query Match      78.8%; Score 93; DB 12; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SCROIKIWFQNRMKWK 20
Db 7 SKROIKIWFQNRMKWK 24

RESULT 20
US-10-651-723-332
; Sequence 332, Application US/10651723
; Publication No. US20040057953A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/651,723
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
```

US-10-651-723-332

Query Match 78.8%; Score 93; DB 12; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 SCRQIKIWFQNRMRMKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 7 SKRQIKIWFQNRMRMKWK 24

RESULT 21

US-10-645-761-332
; Sequence 332, Application US/10645761
; Publication No. US20040071712A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/645,761
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
US-10-645-761-332

Query Match 78.8%; Score 93; DB 12; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 SCRQIKIWFQNRMRMKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 7 SKRQIKIWFQNRMRMKWK 24

RESULT 22

US-10-666-696-332
; Sequence 332, Application US/10666696
; Publication No. US2004007022A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; APPLICANT: GUDAS, JEAN MARIE
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527A
; CURRENT APPLICATION NUMBER: US/10/666,696
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US/09/563,286C
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 09/428,082
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1157
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
US-10-666-696-332

Query Match 78.8%; Score 93; DB 16; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 SCRQIKIWFQNRMRMKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 7 SKRQIKIWFQNRMRMKWK 24

RESULT 23

US-10-653-048-332
; Sequence 332, Application US/10653048
; Publication No. US20040087778A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/653,048
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: P16-MIMETIC
US-10-653-048-332

Query Match 78.8%; Score 93; DB 16; Length 24;
Best Local Similarity 94.4%; Pred. No. 8.8e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 SCRQIKIWFQNRMRMKWK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 7 SKRQIKIWFQNRMRMKWK 24

RESULT 24

US-08-610-220A-9
; Sequence 9, Application US/08610220A
; Publication No. US20030099638A1
; GENERAL INFORMATION:
; APPLICANT: TROY, CAROL M.
; TITLE OF INVENTION: COMPOUNDS WHICH PREVENT NEURONAL CELL
; TITLE OF INVENTION: DEATH AND USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/610,220A
; FILING DATE: MAR-04-1996

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 48332/JPW/JML
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-610-220A-9

Query Match 78.0%; Score 92; DB 8; Length 16;
Best Local Similarity 100.0%; Pred.No. 8.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQIKWQNRRMKWKK 20
Db 1 RQIKWQNRRMKWKK 16

RESULT 25
US-09-214-371-43
Sequence 43, Application US/09214371B
Patent No. US20010018511A1
GENERAL INFORMATION:
APPLICANT: Lane, David
APPLICANT: Bottger, Volker
APPLICANT: Bottger, Angelica
APPLICANT: Pinksley, Stephen
APPLICANT: Chene, Patrick
APPLICANT: Hochkeppel, Heinz-Kurt
APPLICANT: Garcia-Echeverria, Carlos
APPLICANT: Furet, Pascal
TITLE OF INVENTION: Inhibitors of the Interaction of P53 and MDM2
FILE REFERENCE: 4-20937/A/PCT
CURRENT APPLICATION NUMBER: US/09/214,371B
CURRENT FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: PCT/EP97/03549
PRIOR FILING DATE: 1997-07-04
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide
US-09-214-371-43

Query Match 78.0%; Score 92; DB 9; Length 16;
Best Local Similarity 100.0%; Pred.No. 8.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQIKWQNRRMKWKK 20
Db 1 RQIKWQNRRMKWKK 16

Search completed: May 24, 2004, 17:28:44
Job time : 37.5676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:13:28 ; Search time 56.1892 Seconds
(without alignments)
105.599 Million cell updates/sec

Title: US-09-977-349-4
Perfect score: 129
Sequence: 1 CSSCTRQPKIWFNRRKPKWK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A Geneseq 29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	86.8	19	3	AAY93357 Peptide u
2	103	79.8	17	3	AAY93355 Peptide u
3	99	76.7	16	2	AAB33412 Peptide 4
4	99	76.7	16	5	AB83154 Transduct
5	99	76.7	16	6	AAE33893 Drosophil
6	99	76.7	16	6	ABP70226 Membrane
7	99	76.7	26	4	AAE72140 Anti-alle
8	99	76.7	26	4	AAE72143 Anti-alle
9	99	76.7	27	5	AB83155 Transduct
10	85	65.9	16	6	ABU09583 Cell perm
11	82.5	64.0	17	3	AAE83575 Modified
12	81	62.8	18	3	AAE13432 Synthetic
13	79	61.2	17	4	AAE85051 N-termina
14	75	58.1	16	2	AAW33411 Peptide 4
15	75	58.1	16	3	AAE33574 Modified
16	75	58.1	16	3	AAE51213 Antennape
17	75	58.1	16	6	AAE12484 Membrane
18	75	58.1	16	6	AAE35567 Antennape
19	75	58.1	16	6	ABU09582 Cell perm
20	75	58.1	16	6	AAE33892 Drosophil
21	75	58.1	16	6	AB832913 Pro 50 pe
22	75	58.1	16	6	AAO23274 Antennape
23	75	58.1	17	3	AAE13427 Synthetic
24	75	58.1	18	3	AAE93956 Peptide u
25	75	58.1	19	3	AAE13428 Synthetic

26	75	58.1	20	3	AAE13431	Abb13431 Synthetic
27	75	58.1	21	3	AAE13430	Abb13430 Synthetic
28	75	58.1	22	3	AAE13429	Abb13429 Synthetic
29	75	58.1	27	4	AAE12483	Abb12483 Membrane
30	75	58.1	35	6	AAE35578	AAE35578 TA peptid
31	74	57.4	16	4	AAE85059	AAE85059 N-termina
32	74	57.4	36	6	ADA38245	Ada38245 Mutant pe
33	73	56.6	18	3	AAE13425	AAE13425 Synthetic
34	73	56.6	36	6	ADA38244	Ada38244 Peptide p
35	71	55.0	36	6	ABM32769	ABM32769 DGI-3 spe
36	70	54.3	22	2	AAW27475	AAW27475 Human TSP
37	69	53.5	17	3	AAE83558	AAE83558 Peptide f
38	68	52.7	17	5	AAU77232	AAU77232 Synthetic
39	68	52.7	26	3	AAE30393	AAE30393 Beta-cate
40	68	52.7	36	3	AAE71013	AAE71013 Penetrati
41	68	52.7	36	3	AAE71011	AAE71011 Penetrati
42	68	52.7	106	3	AAE21030	AAE21030 Human nuc
43	68	52.7	115	3	AAE53629	AAE53629 Human col
44	68	52.7	220	4	AAW41487	AAW41487 Human pol
45	68	52.7	236	4	AAE39701	AAE39701 Human pol
46	68	52.7	243	2	AAW48885	AAW48885 Amino aci
47	67	51.9	16	2	AAW45974	AAW45974 Cysteine
48	67	51.9	16	2	AAW33407	AAW33407 Peptide 4
49	67	51.9	16	2	AAW33410	AAW33410 D-form pe
50	67	51.9	16	2	AAW82958	AAW82958 Oestrogen
51	67	51.9	16	2	AAW56397	AAW56397 Preferred
52	67	51.9	16	2	AAW71270	AAW71270 Antennape
53	67	51.9	16	2	AAW71316	AAW71316 Antennape
54	67	51.9	16	2	AAW30508	AAW30508 Drosophil
55	67	51.9	16	2	AAW91046	AAW91046 Internal
56	67	51.9	16	2	AAE52102	AAE52102 Peptide f
57	67	51.9	16	2	AAE00859	AAE00859 Peptide p
58	67	51.9	16	2	AAE13509	AAE13509 Signal se
59	67	51.9	16	3	AAE87920	AAE87920 Drosophil
60	67	51.9	16	3	AAE27060	AAE27060 Beta-cate
61	67	51.9	16	3	AAE93667	AAE93667 Peptide w
62	67	51.9	16	3	AAE67966	AAE67966 Carboxyl
63	67	51.9	16	3	AAE93551	AAE93551 Amino aci
64	67	51.9	16	3	AAE55818	AAE55818 Signal se
65	67	51.9	16	3	AAE71008	AAE71008 Drosophil
66	67	51.9	16	3	AAE83572	AAE83572 Modified
67	67	51.9	16	3	AAE51212	AAE51212 Antennape
68	67	51.9	16	3	AAE51167	AAE51167 Drosophil
69	67	51.9	16	3	AAE10343	AAE10343 Peptide A
70	67	51.9	16	3	AAE19251	AAE19251 Fragment
71	67	51.9	16	3	AAE93178	AAE93178 Protegrin
72	67	51.9	16	3	AAE35694	AAE35694 Peptide a
73	67	51.9	16	3	AAE22025	AAE22025 Membrane
74	67	51.9	16	3	AAE29423	AAE29423 ANTP pept
75	67	51.9	16	3	AAE03927	AAE03927 Internal
76	67	51.9	16	3	AAE93954	AAE93954 Peptide u
77	67	51.9	16	3	AAE29574	AAE29574 Antennape
78	67	51.9	16	3	AAE14785	AAE14785 Carrier m
79	67	51.9	16	3	AAE14761	AAE14761 Drosophil
80	67	51.9	16	4	AAE73091	AAE73091 Rheumatoid
81	67	51.9	16	4	AAE60004	AAE60004 Internal
82	67	51.9	16	4	AAE70753	AAE70753 Cell memb
83	67	51.9	16	4	AAE02974	AAE02974 Protein t
84	67	51.9	16	4	AAE60671	AAE60671 Antennape
85	67	51.9	16	4	AAU08064	AAU08064 Drosophil
86	67	51.9	16	4	AAE49914	AAE49914 HIF-1alpha
87	67	51.9	16	4	AAE66996	AAE66996 Antennape
88	67	51.9	16	4	AAU00813	AAU00813 Fruit fly
89	67	51.9	16	4	AAE12205	AAE12205 Membrane
90	67	51.9	16	5	ABE78030	ABE78030 Peptide d
91	67	51.9	16	5	ABE78985	ABE78985 Cell pene
92	67	51.9	16	5	AAW47331	AAW47331 Penetrati
93	67	51.9	16	5	AAE78214	AAE78214 Amino aci
94	67	51.9	16	5	AAE15616	AAE15616 Drosophil
95	67	51.9	16	5	AAO15781	AAO15781 Antennape
96	67	51.9	16	5	AAU78345	AAU78345 Antennape
97	67	51.9	16	5	ABG68406	ABG68406 Translaca
98	67	51.9	16	5	ABE83153	ABE83153 Transduct

99 67 51.9 16 5 ABP53809 Penetrati
100 67 51.9 16 5 AAE25451 Drosophil

ALIGNMENTS

RESULT 1
AA93957
ID AAY93957 standard; peptide; 19 AA.
AC AAY93957;
XX
XX 03-OCT-2000 (first entry)
XX
XX Peptide used to construct inhibitory pseudopeptide compounds.
XX
XX Pseudopeptide compound; inhibitor; tyrosine kinase;
KW growth factor receptor binding protein; Grb2; proliferative disorder;
KW cancer; metastasis; antennapedia protein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Disulfide-bond 1..2
FT
XX WO200039153-A1.
XX
XX 06-JUL-2000.
XX
XX 24-DEC-1999; 99WO-FR003289.
XX
XX 24-DEC-1998; 98FR-00016459.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (CNRS) CNRS CENT NAT RECH SCT.
XX Garbay C, Liu W, Vidal M, Roques BP;
XX
XX WPI; 2000-475683/41.
XX
XX Pseudopeptides, useful for the treatment of proliferative process
PT disorders, cancers, and metastases, inhibit paths activated by proteins
FT with a tyrosine kinase activity.
XX
XX Claim 1; Page 32; 43pp; French.
XX
XX The specification describes pseudopeptide compounds. These compounds are
CC inhibitors of routes activated by proteins having a tyrosine kinase
CC activity. The compounds also have high affinity for a growth factor
CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
CC the treatment of disorders of proliferative processes, cancers, and
CC metastases. The present sequence is derived from the antennapedia
CC protein, and is a fragment of the pseudopeptide compounds of the
CC invention
XX
XX Sequence 19 AA;
XX
Query Match 86.8%; Score 112; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 CIRQPKIWFNNRRKPWK 21
Db 2 CIRQPKIWFNNRRKPWK 19
RESULT 2
AAY93955
ID AAY93955 standard; peptide; 17 AA.
XX
XX AC AAY93955;
XX

DT 03-OCT-2000 (first entry)
XX
XX Peptide used to construct inhibitory pseudopeptide compounds.
XX
XX Pseudopeptide compound; inhibitor; tyrosine kinase;
KW growth factor receptor binding protein; Grb2; proliferative disorder;
KW cancer; metastasis; antennapedia protein.
XX
XX Synthetic.
XX
XX WO200039153-A1.
XX
XX 06-JUL-2000.
XX
XX 24-DEC-1999; 99WO-FR003289.
XX
XX 24-DEC-1998; 98FR-00016459.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (CNRS) CNRS CENT NAT RECH SCT.
XX Garbay C, Liu W, Vidal M, Roques BP;
XX
XX WPI; 2000-475683/41.
XX
XX Pseudopeptides, useful for the treatment of proliferative process
PT disorders, cancers, and metastases, inhibit paths activated by proteins
FT with a tyrosine kinase activity.
XX
XX Claim 1; Page 32; 43pp; French.
XX
XX The specification describes pseudopeptide compounds. These compounds are
CC inhibitors of routes activated by proteins having a tyrosine kinase
CC activity. The compounds also have high affinity for a growth factor
CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
CC the treatment of disorders of proliferative processes, cancers, and
CC metastases. The present sequence is derived from the antennapedia
CC protein, and is a fragment of the pseudopeptide compounds of the
CC invention
XX
XX Sequence 17 AA;
XX
Query Match 79.8%; Score 103; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 5 IROPKIWFNNRRKPWK 21
Db 1 IROPKIWFNNRRKPWK 17
RESULT 3
AAW33412
ID AAW33412 standard; peptide; 16 AA.
XX
XX AAW33412;
XX
XX 17-MAR-1998 (first entry)
XX
XX Peptide 43-58 (3Pro) of homeodomain Antp.
XX homeodomain; transcription factor; Antennapedia; Antp; vector;
KW transfection; hydrophobic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1
FT /note= "in determining the ability of this sequence to be
FT internalised in cells, a biotin-aminopentanoyl group was
FT attached to the N-terminal"
XX
XX WO9712912-A1.
PN

XX PD 10-APR-1997.
 XX PF 04-OCT-1996; 96WO-FR001553.
 XX PR 05-OCT-1995; 95FR-00011714.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Chassaing G, Prochiantz A;
 XX DR WPI; 1997-226166/20.
 XX PT New peptide(s) of high hydrophobic amino acid content - useful as vectors
 XX for delivering peptides and nucleic acids to cells.
 XX PS Disclosure; Page 7; 35pp; French.
 XX CC New peptides are provided which are 16 amino acids long and which are
 CC analogues of the peptide corresponding to residues 43-58 of the
 CC Antennapedia transcription factor homeodomain (AntpHD). The peptides
 CC contain 6-10 hydrophobic amino acids. They have the general formula: X1-
 CC X2-X3-X4-X5-Trp-X7-X8-X9-X10-X12-X13-X14-X15-X16 or X16-X15-X14-X13-
 CC X12-X11-X10-X9-X8-X7-Trp-X5-X4-X3-X2-X1 in which X1-X5 and X7-X16 are any
 CC alpha-amino acids, provided that: (1) the peptide contains 6-10
 CC hydrophobic amino acids; (2) X3 and X5 are not both Val; and (3) the
 CC natural Antp 43-58 sequence RQKWFQNRKRWKK (see AAW33407) is excluded.
 CC The present sequence (a pro analogue of the 43-58 sequence) is a specific
 CC example of the new peptides. The peptides are used as vectors for
 CC introducing into live cells compounds which affect cell function,
 CC especially peptides and nucleic acids. They can cross cellular membranes
 CC and reach various cell compartments. They are as effective as helix 3 of
 CC a homeodomain peptide
 XX SQ Sequence 16 AA;
 Query Match 76.7%; Score 99; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 RQKWFQNRKRWKK 21
 Db 1 RQKWFQNRKRWKK 16
 |||||
 RESULT 4
 ABB83154
 ID ABB83154 standard; peptide; 16 AA.
 AC ABB83154;
 XX DT 08-AUG-2002 (first entry)
 DE Transduction domain of pAntp peptide, Pen3P.
 XX KW Transduction domain; nuclear export; tight epithelial junction;
 XX KW blood-brain barrier; choroid plexus; antennapedia.
 XX OS Drosophila sp.
 XX PN WO200239947-A2.
 XX PD 23-MAY-2002.
 XX PF 20-NOV-2001; 2001WO-FR003631.
 XX PR 20-NOV-2000; 2000FR-00014945.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Joliot A, Dupont E, Prochiantz A;
 XX DR WPI; 2002-471597/50.

XX PT Use of vector a peptide containing transduction domain and nuclear export
 PT sequence for transporting diagnostic or therapeutic agents across tight
 PT epithelial junctions.
 XX PS Example 4; Page 16; 30pp; French.
 XX CC The present invention relates to peptides that include at least one
 CC transduction domain (TD) and a nuclear export sequence (NES) to prepare a
 CC vector for transporting diagnostic or therapeutic agents across a tight
 CC epithelial junction. The present sequence is one such TD, the TD from the
 CC third helix of the pAntp peptide of the Antennapedia protein from
 CC Drosophila. The combination of TD and NES increases the efficiency of
 CC transport through tight junctions, compared with use of TD alone. The
 CC vectors are useful for delivering nucleic acids, polypeptides, peptide
 CC nucleic acids, or a nucleotide analog (e.g. an antineoplastic agent)
 CC across the blood-brain barrier or the choroid plexus
 XX SQ Sequence 16 AA;
 Query Match 76.7%; Score 99; DB 5; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 RQKWFQNRKRWKK 21
 Db 1 RQKWFQNRKRWKK 16
 |||||
 RESULT 5
 AAE33893
 ID AAE33893 standard; peptide; 16 AA.
 XX AC AAE33893;
 XX DT 02-MAY-2003 (first entry)
 XX DE Drosophila sp. antennapedia third helix peptide #5.
 XX KW RNA polymerase; cancer; autoimmune disease; cardiovascular disease;
 KW arthritis; cystic fibrosis; infection; sickle cell anaemia; herpes; AIDS;
 KW pneumonia; tuberculosis; inflammatory disease; gene therapy; cytostatic;
 KW immunosuppressive; cardiac; virucide; antiinflammatory.
 XX OS Drosophila sp.
 XX PN WO200288370-A2.
 XX PD 07-NOV-2002.
 XX PF 30-APR-2002; 2002WO-CA000670.
 XX PR 30-APR-2001; 2001US-0287974P.
 XX PA (PROT-) PROTIVA BIOTHERAPEUTICS INC.
 XX PI Finn J, Maclachlan I;
 XX DR WPI; 2003-156691/15.
 XX PT New nucleic acid for treating diseases, e.g. cancer, autoimmune disease,
 PT cardiovascular disease or AIDS, comprises a secretable RNA polymerase
 PT expression cassette having a eukaryotic promoter and an RNA polymerase
 PT promoter.
 XX PS Claim 5; Page 4; 65pp; English.
 XX CC The invention relates to autogene nucleic acids encoding secretable RNA
 CC polymerases. The invention also relates to methods, nucleic acids,
 CC compounds and compositions for expressing a product of interest in a cell
 CC that involve a secretable RNA polymerase. The nucleic acid is useful in
 CC treating a disease, e.g. cancer, an autoimmune disease, arthritis, a
 CC cardiovascular disease, cystic fibrosis, an infectious disease, sickle

CC cell anaemia, a viral disease, AIDS, a bacterial disease, herpes,
CC pneumonia, tuberculosis or an inflammatory disease, in a patient. It is
CC also used in gene therapy. The present sequence is Drosophila sp.
CC antennapedia third helix peptide used to illustrate the method of the
CC invention
XX
SQ Sequence 16 AA;

Query Match 76.7%; Score 99; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RQPKIWFNRRKPWK 21
Db 1 RQPKIWFNRRKPWK 16

RESULT 6
ABP70226
ID ABP70226 standard; peptide; 16 AA.
XX
AC ABP70226;
XX
DT 07-APR-2003 (first entry)
XX
DE Membrane translocating peptide from homeobox domain penetration region.
XX
KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;
KW cancer; infection; immune deficiency; gene defect; genetic disease;
KW membrane translocating peptide.
XX

OS Unidentified.
XX
PN WO200288318-A2.
XX
PD 07-NOV-2002.
XX
PF 30-APR-2002; 2002WO-US013609.
XX
PR 30-APR-2001; 2001US-0287786P.
XX
PA (TARG-) TARGETED GENETICS CORP.
PA (EMER-) EMERALD GENE SYSTEMS LTD.
XX
PI Harvie P, Paul R, Cudmore S, O'mahony DJ;
XX
DR WPI; 2003-183837/18.

XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a
PT cell, comprises compacted nucleic acid, polycation, targeting factor and
PT lipid, and does not comprise protamine or its salt.
XX

PS Disclosure; Page 42; 259pp; English.
XX
XX The specification describes a lipid-nucleic acid complex, comprising a
CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but
CC not a protamine. The targeting factor increases cellular bioavailability
CC of the nucleic acid without interaction with a specific outer cell
CC surface membrane receptor. The mean diameter of the complex is greater
CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is
CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene
CC therapy. It reduces levels of inflammatory cytokines such as tumour
CC necrosis factor-alpha. The complex is useful for manufacturing a
CC medicament for treating or diagnosing a variety of diseases, conditions
CC or syndromes such as cancer, bacterial, viral or parasitic infections,
CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited
CC genetic diseases). The present sequence represents a membrane
CC translocating peptide, which is used as the targeting factor in lipid-
CC nucleic acid complexes of the invention

XX Sequence 16 AA;

Query Match 76.7%; Score 99; DB 6; Length 16;

Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 RQPKIWFNRRKPWK 21
Db 1 RQPKIWFNRRKPWK 16

RESULT 7
AAY72140
ID AAY72140 standard; peptide; 26 AA.
XX
AC AAY72140;

XX
DT 24-APR-2001 (first entry)
XX
DE Anti-allergic peptide 3.

XX
KW Fruit fly; anti-allergic peptide; therapeutic; migraine; psoriasis;
KW multiple sclerosis; nasal allergy; mast cell degranulation; histamine;
KW allergy; eye; skin; acute urticaria; interstitial cystitis; vasotropic;
KW psychogenic; bowel disease; dermatological; antiinflammatory; G alphas3;
KW neuroprotective; antipsoriatic; fusion peptide; exocytosis; asthma;
KW Drosophila transcription factor.

XX
OS Drosophila sp.
OS Unidentified.
OS Chimeric.

XX
FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= Signal peptide
FT /note= "Signal sequence of homeodomain of Drosophila
FT transcription factor"
FT Peptide 17..26
FT /label= G alphas3_peptide
FT /note= "Corresponds to C-terminal sequence of G alphas3"

XX WO200078346-A1.

XX 28-DEC-2000.

XX 14-JUN-2000; 2000WO-IL000346.

XX 17-JUN-1999; 99IL-00130526.

XX (ALB-) ALLERGENE LTD.

XX Eisenberg R, Raz T;

XX WPI; 2001-080758/09.

XX Novel anti-allergic agents for treating allergic conditions such as
PT allergic reactions in eye, skin, nasal allergy, asthma, migraines, has
PT peptides for cell penetration and reducing mast cell degranulation.

XX Example 1; Page 13; 63pp; English.

XX The present sequence is anti-allergic peptide 3 consisting of a signal
CC sequence of homeodomain of Drosophila transcription factor, linked to the
CC C-terminal G alphas3 sequence. This C-terminal G alphas3 appears to
CC mediate the peptidergic pathway leading to exocytosis in mast cells. The
CC invention relates to therapeutic complex molecules which are useful as
CC anti-allergic agents. These anti-allergic agents are useful for treating
CC allergic conditions such as nasal allergy, allergic reaction in the eye
CC or skin, acute urticaria, psoriasis, psychogenic or allergic asthma,
CC interstitial cystitis, bowel diseases, migraines and multiple sclerosis.
CC The therapeutic complex is highly specific, direct and provides targeted
CC treatment of allergies and related inflammatory conditions. It comprises
CC molecules having at least a first segment ie., a signal peptide which is
CC competent for the importation of the complex into the mast cells, and a
CC second segment which is having the anti-allergic effect is able to block
CC or significantly reduce the G protein-mediated contribution to mast cell

CC degranulation and in turn the release of histamine. The invention also
CC discloses methods for preventing and treating allergies

XX
SQ Sequence 26 AA;

Query Match 76.7%; Score 99; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. NO. 3.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKIWFNNRRKPWK 21
DB 1 RQPKIWFNNRRKPWK 16

RESULT 8
AAV72143
ID AAV72143 standard; peptide; 26 AA.
XX
AC AAV72143;

DT 24-APR-2001 (first entry)
XX
DE Anti-allergic peptide 6.

KW Fruit fly; anti-allergic peptide; therapeutic; migraine; psoriasis;
KW multiple sclerosis; nasal allergy; mast cell degranulation; histamine;
KW allergy; eye; skin; acute urticaria; interstitial cystitis; vasotropic;
KW psychogenic; bowel disease; dermatological; antiinflammatory; G alphet;
KW neuroprotective; antipsoriatic; fusion peptide; exocytosis; asthma;
KW Drosophila transcription factor.

XX
OS Drosophila sp.
OS Unidentified.
OS Chimeric.

XX
FH Key Location/Qualifiers

FT Peptide 1..16
FT /label= Signal peptide
FT /note= "Signal sequence of homeodomain of Drosophila

FT transcription factor"
FT 17..26
FT /label= G alphet peptide
FT /note= "Corresponds to C-terminal sequence of G alphet"

XX WO200078346-A1.

XX 28-DEC-2000.

XX 14-JUN-2000; 2000WO-IL000346.

XX 17-JUN-1999; 99IL-00130526.

XX (ALLE-) ALLERGENE LTD.

XX Eisenberg R, Raz T;

XX WPI; 2001-080758/09.

XX Novel anti-allergic agents for treating allergic conditions such as
XX allergic reactions in eye, skin, nasal allergy, asthma, migraines, has
XX peptides for cell penetration and reducing mast cell degranulation.

PS Example 1; Page 13; 63pp; English.

XX The present sequence is anti-allergic peptide 6 consisting of a signal
XX sequence of homeodomain of Drosophila transcription factor, linked to the
XX C-terminal G alphet sequence. The invention relates to therapeutic
XX complex molecules which are useful as anti-allergic agents. These anti-
XX allergic agents are useful for treating allergic conditions such as nasal
XX allergy, allergic reaction in the eye or skin, acute urticaria,
XX psoriasis, psychogenic or allergic asthma, interstitial cystitis, bowel
XX diseases, migraines and multiple sclerosis. The therapeutic complex is
XX highly specific, direct and provides targeted treatment of allergies and

CC related inflammatory conditions. It comprises molecules having at least a
CC first segment ie., a signal peptide which is competent for the
CC importation of the complex into the mast cells, and a second segment
CC which is having the anti-allergic effect is able to block or
CC significantly reduce the G protein-mediated contribution to mast cell
CC degranulation and in turn the release of histamine. The invention also
CC discloses methods for preventing and treating allergies

XX
SQ Sequence 26 AA;

Query Match 76.7%; Score 99; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. NO. 3.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RQPKIWFNNRRKPWK 21
DB 1 RQPKIWFNNRRKPWK 16

RESULT 9
ABB83155
ID ABB83155 standard; peptide; 27 AA.
XX
AC ABB83155;

DT 08-AUG-2002 (first entry)

DE Transduction domain/nuclear export sequence peptide NES-Pen3P.

XX Transduction domain; nuclear export; tight epithelial junction;

KW blood-brain barrier; choroid plexus.

XX Unidentified.

XX
FH Key Location/Qualifiers

FT Modified-site 1 /note= "Biotinylated Gln"

XX WO200239947-A2.

XX 23-MAY-2002.

XX 20-NOV-2001; 2001WO-FR003631.

XX 20-NOV-2000; 2000FR-00014945.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Joliot A, Dupont E, Prochiantz A;

XX WPI; 2002-471597/50.

XX Use of vector a peptide containing transduction domain and nuclear export
XX sequence for transporting diagnostic or therapeutic agents across tight
XX epithelial junctions.

PS Example 4; Page 16; 30pp; French.

XX The present invention relates to peptides that include at least one
XX transduction domain (TD) and a nuclear export sequence (NES) to prepare a
XX vector for transporting diagnostic or therapeutic agents across a tight
XX epithelial junction. The present sequence is one such peptide. The
XX combination of TD and NES increases the efficiency of transport through
XX tight junctions, compared with use of TD alone. The vectors are useful
XX for delivering nucleic acids, polypeptides, peptide nucleic acids, or a
XX nucleotide analog (e.g. an antineoplastic agent) across the blood-brain
XX barrier or the choroid plexus

XX
SQ Sequence 27 AA;

Query Match 76.7%; Score 99; DB 5; Length 27;
Best Local Similarity 100.0%; Pred. NO. 3.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRKPKWK 21
 DB 12 ROPKIWFPPNRKPKWK 27

RESULT 10
 ABU09583
 ID ABU09583 standard; peptide; 16 AA.

XX AC ABU09583;

DT 09-JUL-2003 (first entry)

XX DE Cell permeant peptide #4.

XX KW Capsid protein; CCD; adaptin library; cell permeant peptide; cell death;
 KW pathogen; viral infection; bacterial infection; toxin; Rickets;
 KW fungal infection; mortality; morbidity; biological warfare.

XX OS Unidentified.

XX PN US2002192799-A1.

XX PD 19-DEC-2002.

XX PF 15-OCT-2001; 2001US-00981286.

XX PR 13-OCT-2000; 2000US-0240187P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Watowich SJ, Weaver SC, Davey RA;

XX PR WPI; 2003-428794/40.

XX PT Novel collection of polypeptides which comprise fragment of Venezuelan
 PT equine encephalitis virus capsid polypeptide carboxy terminal portions,
 PT useful for displaying variable amino acid sequences on the surface of the
 PT polypeptide.

XX PS Claim 3; Page 12; 25pp; English.

XX CC The invention relates to a collection of polypeptides comprising two
 CC polypeptides (referred to as an adaptin library) each of which has a
 CC fragment of amino acids 119-275 of Venezuelan equine encephalitis virus
 CC capsid polypeptide carboxy terminal portion (ABU09576, also known as
 CC CCD), beginning at any of amino acids 119-124 and ending at any of amino
 CC acids 258-275, where at least two consecutive amino acids within amino
 CC acids 129-137, 182-189, 257-264 of CCD are replaced by a variant amino
 CC acid sequence. Also included are a population of cells comprising two or
 CC more cells (where each member of the population comprises one polypeptide
 CC of an adaptin), a fusion protein chosen from an adaptin fused to a cell
 CC permeant peptide, a cell comprising the fusion protein, a collection of
 CC polynucleotides (comprising at least two polynucleotides, each
 CC polynucleotide comprising a coding sequence encoding a polypeptide
 CC comprising an adaptin), a vector comprising a member of the
 CC polynucleotide collection, a population of cells comprising two or more
 CC cells containing members of the collection of fusion proteins and
 CC crystallising a polypeptide comprising CCD. The adaptins are useful for
 CC identifying a polypeptide within a collection that prevents cell death
 CC after exposure to a pathogen (e.g. a virus or microbe such as bacterium,
 CC Rickettsia or fungus), or a toxin such as a biological toxin or chemical
 CC toxin and is useful for identifying a polypeptide within a collection
 CC that binds to a pathogen, toxin, polypeptide or a polynucleotide. The
 CC adaptin is useful in discovery of drugs that can be used to prevent
 CC mortality and morbidity associated with pathogens and toxins, including
 CC high lethal agents that could be used in biological warfare. The present
 CC sequence represents a cell permeant peptide used to make fusion proteins
 CC with the CCD-based adaptins

XX SQ Sequence 16 AA;

Query Match 65.9%; Score 85; DB 6; Length 16;
 Best Local Similarity 87.5%; Pred. No. 2.4e-05;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ROPKIWFPPNRKPKWK 21
 DB 1 ROPKIWFPPNRKPKWK 16

RESULT 11

AA83575
 ID AAY83575 standard; peptide; 17 AA.

XX AC AAY83575;

XX DT 29-AUG-2000 (first entry)

XX DE Modified peptide 45,50,55 Pro of penetratin.
 XX KW Penetratin; translocation; membrane; drug delivery; therapy; treatment;

XX KW conjugate; vector.

XX OS Synthetic.

XX PN WO200029427-A2.

XX PD 25-MAY-2000.

XX PF 11-NOV-1999; 99WO-GB003750.

XX PR 13-NOV-1998; 98GB-00025000.

XX PR 04-FEB-1999; 99GB-00002522.

XX PR 04-FEB-1999; 99GB-00002525.

XX PR 22-JUN-1999; 99GB-00014578.

XX PA (CYCL-) CYCLACEL LTD.

XX PI Fischer MP, Zhelev N;

XX DR WPI; 2000-387734/33.

XX PT New membrane translocation peptide carrier group for delivering
 PT therapeutic agents into target cells comprises specified sequence of
 PT amino acids.

XX PS Example 4; Page 29; 59pp; English.

XX CC Penetratin is a membrane translocation polypeptide and as such, active
 CC peptide fragments of penetratin can be used to translocate conjugated
 CC molecules across the cell membrane e.g. drugs. 13 chemical entities are
 CC described for use in the method such as paclitaxel-
 CC 2'succinimidopropionyl-Cbetaa-RMKWK-NH₂ and podophyllotoxin-4-
 CC succinimidopropionyl Cbetaa-RMKWK-NH₂. The method has applications as
 CC a drug delivery system for treatment and therapy. The resulting
 CC conjugated molecules exhibit high immunogenicity, solubility and
 CC clearance. The penetratin peptide fragment may be truncated and or have
 CC amino acid substitutions. See GENESEQ records AAY83520-Y83581

XX SQ Sequence 17 AA;

Query Match 64.0%; Score 82.5; DB 3; Length 17;
 Best Local Similarity 88.2%; Pred. No. 5.7e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 6 ROPKIWFPPNRKPKWK 21
 DB 1 ROPKIWFPPNRKPKWK 17

RESULT 12

AA813432
 ID AA813432 standard; peptide; 18 AA.

XX AC AAB13432;
 XX DT 23-NOV-2000 (first entry)
 XX DE Synthetic alpha smooth muscle actin inhibitor # 12.
 XX KW Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX OS Synthetic.
 XX FH Key
 FT Cross_link 1 Location/Qualifiers
 FT /note= "Pentapeptide of Acetyl-BEEDC is linked via a
 FT disulphide bond to the thiol group of the side chain of
 FT the cysteine residue"
 XX WO200038733-A1.
 XX PN 06-JUL-2000.
 XX PD 15-DEC-1999; 99WO-EF009964.
 XX PF 24-DEC-1998; 98EP-00204396.
 XX PR (UNIO) UCB-BIOPRODUCTS SA.
 XX PA Gabbiani G, Scarso A;
 XX PI WPI; 2000-452308/39.
 XX DR A peptidic product for prevention and treatment of a disease related to
 XX PT alpha-SM actin expression comprises a tetrapeptide associated with a
 XX PT chemical entity that is able to introduce the tetrapeptide into the cell.
 XX PS Disclosure; Page; 31pp; English.
 XX CC The present invention relates to novel peptides comprising of a specific
 CC tetrapeptide associated with an oligopeptide which allows the
 CC introduction of the tetrapeptide into the target cell. The present
 CC sequence is one such peptide. The side chain linked to the lysine residue
 CC at position 1 corresponds to the specific tetrapeptide, while residues 1
 CC to 17 correspond to the oligopeptide. The specific tetrapeptide of the
 CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
 CC organisation in stress fibers. The present sequence may be used in the
 CC prevention and/or treatment of a disease related to alpha-SM actin
 CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
 CC fibrotic conditions. The present sequence may also be used to treat
 CC Dupuytren disease and lung fibrosis. Note: the present sequence is not
 CC shown in the specification but is derived from information given by the
 CC inventors
 XX CC
 XX SQ Sequence 18 AA;
 Query Match 62.8%; Score 81; DB 3; Length 18;
 Best Local Similarity 77.8%; Pred. No. 0.0001;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 4 CIRQPKIWFNRRKPKK 21
 ||| ||||| |||||
 Db 1 CKRQIKIWFNRRKPKK 18
 RESULT 13
 AAB85051
 ID AAB85051 standard; peptide; 17 AA.
 XX AC AAB85051;
 XX KW AAB85051;
 XX OS Synthetic.
 XX DT 06-AUG-2001 (first entry)
 XX FH Key
 FT Location/Qualifiers

DE N-terminal S-benzyl thioester peptide #1.
 XX Coupling; amine derivative; hydroxy derivative; thio ester; thio ether.
 XX OS Synthetic.
 XX FH Key
 FT Modified-site 1 Location/Qualifiers
 FT /note= "attached to BnSCO"
 FT 5..6
 FT Modified-site /note= "attached to Tfa"
 FT 13..14
 FT Modified-site /note= "attached to Tfa"
 FT 16..17
 FT Modified-site /note= "attached to Tfa"
 FT 17
 FT Modified-site /note= "C-terminal CONH2"
 FT 17
 FT /note= "attached to Tfa"
 XX WO200115737-A2.
 XX PN 08-MAR-2001.
 XX PD 25-AUG-2000; 2000WO-GB003306.
 XX PF 27-AUG-1999; 99GB-00020397.
 XX PR 18-MAY-2000; 2000GB-00012083.
 XX PA (MEDI-) MEDICAL RES COUNCIL.
 XX PI Gait MJ, Stetsenko D;
 XX WPI; 2001-367105/38.
 XX DR Method for coupling molecules, e.g. peptides and oligonucleotides, and
 XX PT new intermediates and reagents.
 XX PS Example 3; Page 27; 39pp; English.
 XX CC The invention relates to methods for coupling molecules. One method of
 CC linking an amine derivative with a hydroxy derivative comprises reacting
 CC an alkyl or aryl thio ester with an amino thioether. The compounds of the
 CC invention are useful e.g. for linking a peptide and an oligonucleotide.
 CC Sequences AAB85051-59 represent peptide N-terminal S-benzyl thio esters
 XX SQ Sequence 17 AA;
 Query Match 61.2%; Score 79; DB 4; Length 17;
 Best Local Similarity 92.9%; Pred. No. 0.00018;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 8 PKIWFNRRKPKK 21
 ||| ||||| |||||
 Db 4 PKIWFNRRKPKK 17
 RESULT 14
 AAW33411
 ID AAW33411 standard; peptide; 16 AA.
 XX AC AAW33411;
 XX DT 17-MAR-1998 (first entry)
 XX DE Peptide 43-58 (Pro50) of homeodomain Antp.
 XX KW homeodomain; transcription factor; Antennapedia; Antp; vector;
 KW transfection; hydrophobic.
 XX OS Synthetic.
 XX FH Key
 FT Location/Qualifiers

EC polynucleotide collection, a population of cells comprising two or more

CC anemnapetura curru nexix peptide used to illustrate the method of the
CC invention

XX SQ Sequence 16 AA;

Query Match 58.1%; Score 75; DB 6; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.00065;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPWK 21
||| ||||| |||
Db 1 RQKIWFPPNRRKPWK 16

RESULT 21

ABB82913
ID ABB82913 standard; peptide; 16 AA.

XX AC ABB82913;

DT 31-MAR-2003 (first entry)

XX DE Pro 50 peptide fragment.

XX KW Growth factor; interleukin; antioxidant; collagen; pharmaceutical;
XX KW cosmetic; transport peptide; Pro 50.

XX OS Unidentified.

XX PN WO200298365-A2.

XX PD 12-DEC-2002.

XX PF 07-JUN-2002; 2002WO-US018057.

XX PR 07-JUN-2001; 2001US-0297177P.

XX PA (ADTI-) ADVANCED TISSUE SCI INC.

XX PI Mansbridge J;

XX DR WPI; 2003-140541/13.

XX CC Composition comprising conditioned cell culture media which comprises a culture-derived growth factor (e.g. vascular endothelial growth factor), an antioxidant (e.g. Glutathione), and soluble collagen.

XX PS Claim 21; Page 16; 74pp; English.

XX CC The invention provides a composition comprising, conditioned cell culture media, or its extract, comprising at least one culture-derived growth factor such as vascular endothelial growth factor (VEGF), transforming growth factor beta (TGFbeta), hepatocyte growth factor (HGF), keratinocyte growth factor (KGF), interleukin-3 (IL-3), IL-6 or IL-8, at least one culture-derived antioxidant such as glutathione, glutathione peroxidase, glutathione reductase, glutathione disulfide, catalase, superoxide dismutase, alpha-tocopherol, gamma-tocopherol, ubiquinol-9, ubiquinone 9, ascorbic acid, cysteine and cystine, and at least one culture-derived soluble collagen, and an appropriate carrier. The composition is useful in cosmetic applications, pharmaceutical applications, pharmaceutical applications etc. Sequences ABB82912-930 represent exemplary transport peptides known to enhance cell membrane permeation or transport and forms a part of the composition of the invention

XX SQ Sequence 16 AA;

Query Match 58.1%; Score 75; DB 6; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.00065;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPWK 21
||| ||||| |||
Db 1 RQKIWFPPNRRKPWK 16

RESULT 22

AAO23274

ID AAO23274 standard; peptide; 16 AA.

XX AC AAO23274;

DT 25-SEP-2003 (first entry)

XX DE Antennapedia homeodomain peptide to target intracellular polypeptides.

XX KW Immunoregulation; T cell response; 4-LBB binding agent; CD137; TNF;
XX KW tumour necrosis factor; immunogenic; gene therapy; leukaemia; cancer;
XX KW astrocytoma; infectious disease; immunostimulant; antennapedia; anergy.

XX OS Unidentified.

XX PN WO2003049755-A1.

XX PD 19-JUN-2003.

XX PF 09-OCT-2002; 2002WO-US032364.

XX PR 09-OCT-2001; 2001US-0328004P.

XX PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

XX PI Chen L, Strome SE;

XX DR WPI; 2003-532879/50.

XX CC Generating an enhanced immune response in a subject comprises administering to the subject an immunogenic stimulus (e.g. a tumor-associated antigen) and an agonistic 4-LBB-binding agent (e.g. antibody that binds to 4-LBB).

XX PS Disclosure; Page 21; 81pp; English.

XX CC This invention relates to a novel method of immunoregulation, specifically T cell response regulation by enhancing the immune response with 4-LBB binding agents. 4-LBB (also known as CD137) is a member of the tumour necrosis factor (TNF) receptor family, and is expressed on the surface of activated T-cells. Administration of a 4-LBB binding agent and a tumour specific immunogenic stimulus has been found to result in the regression of that tumour. The binding agent can be an antibody that binds 4-LBB or a natural ligand of 4-LBB (or a functional fragment thereof), whereas the immunogenic stimulus is either a tumour associated peptide fragment or a molecule produced by an infectious microorganism e.g. a retrovirus or protozoan parasite. As such, the present invention describes a method of gene therapy for the treatment of, for example, leukaemia, astrocytomas, colon cancer, testicular cancer, bone and vascular tumours. In addition, with an appropriate immunostimulant it can provide a therapy or prophylaxis for various infectious diseases. Furthermore, it has been shown that this method prevents anergy (a depression in immunological function), and can also reverse established anergy, in CD8+ T-cells. This peptide sequence is the antennapedia homeodomain peptide that acts to direct proteins, such as the immunostimulant or 4-LBB binding agents, to specific cellular compartments, the method of the invention

XX SQ Sequence 16 AA;

Query Match 58.1%; Score 75; DB 6; Length 16;
Best Local Similarity 81.2%; Pred. No. 0.00065;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 RQKIWFPPNRRKPWK 21
||| ||||| |||
Db 1 RQKIWFPPNRRKPWK 16

RESULT 23

AAB13427

ID AAB13427 standard; peptide; 17 AA.
 XX
 AC AAB13427;
 XX
 DT 23-NOV-2000 (first entry)
 XX
 DE Synthetic alpha smooth muscle actin inhibitor # 7.
 XX
 KW Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX
 OS Synthetic.
 XX
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Tetrapeptide of Acetyl-EEED linked on the epsilon
 FT -amino group of lysyl residue"
 XX
 PN WO200038733-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 15-DEC-1999; 99WO-EP009964.
 XX
 PR 24-DEC-1998; 98EP-00204396.
 XX
 PA (UNIO) UCB-BIOPRODUCTS SA.
 XX
 PI Gabbiani G, Scarso A;
 XX
 DR WPI; 2000-452308/39.
 XX
 XX A peptidic product for prevention and treatment of a disease related to
 PT alpha-SM actin expression comprises a tetrapeptide associated with a
 PT chemical entity that is able to introduce the tetrapeptide into the cell.
 XX
 PS Claim 7; Page 24; 31pp; English.
 XX
 CC The present invention relates to novel peptides comprising of a specific
 CC tetrapeptide associated with an oligopeptide which allows the
 CC introduction of the tetrapeptide into the target cell. The present
 CC sequence is one such peptide. The side chain linked to the lysine residue
 CC at position 1 corresponds to the specific tetrapeptide, while residues 1
 CC to 17 correspond to the oligopeptide. The specific tetrapeptide of the
 CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
 CC organisation in stress fibers. The present sequence may be used in the
 CC prevention and/or treatment of a disease related to alpha-SM actin
 CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
 CC fibrotic conditions. The present sequence may also be used to treat
 CC Dupuytren disease and lung fibrosis. The tetrapeptide of the present
 CC invention is described in AAB14263
 XX
 SQ Sequence 17 AA;
 Query Match 58.1%; Score 75; DB 3; Length 17;
 Best Local Similarity 81.2%; Pred. No. 0.00069;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 6 RQPKIWFNRRKPWK 21
 DB 2 RQIKIWFNRRKPWK 17
 RESULT 24
 AAY93956
 ID AAY93956 standard; peptide; 18 AA.
 XX
 AC AAY93956;
 XX
 DT 03-OCT-2000 (first entry)
 XX
 DE Peptide used to construct inhibitory pseudopeptide compounds.

XX Pseudopeptide compound; inhibitor; tyrosine kinase;
 KW growth factor receptor binding protein; Grb2; proliferative disorder;
 KW cancer; metastasis; antennapedia protein.
 XX
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT Disulfide-bond 1..2
 XX WO200039153-A1.
 PN
 PD 06-JUL-2000.
 XX
 PF 24-DEC-1999; 99WO-FR003289.
 XX
 PR 24-DEC-1998; 98FR-00016459.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Garbay C, Liu W, Vidal M, Roques BP;
 XX
 DR WPI; 2000-475683/41.
 XX
 XX Pseudopeptides, useful for the treatment of proliferative process
 PT disorders, cancers, and metastases, inhibit paths activated by proteins
 PT with a tyrosine kinase activity.
 XX
 PS Claim 1; Page 32; 43pp; French.
 XX
 CC The specification describes pseudopeptide compounds. These compounds are
 CC inhibitors of routes activated by proteins having a tyrosine kinase
 CC activity. The compounds also have high affinity for a growth factor
 CC receptor binding protein (Grb2). The pseudopeptide compounds are used for
 CC the treatment of disorders of proliferative processes, cancers, and
 CC metastases. The present sequence is derived from the antennapedia
 CC protein, and is a fragment of the pseudopeptide compounds of the
 CC invention
 XX
 SQ Sequence 18 AA;
 Query Match 58.1%; Score 75; DB 3; Length 18;
 Best Local Similarity 72.2%; Pred. No. 0.00073;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 CIRQPKIWFNRRKPWK 21
 DB 1 CCRQIKIWFNRRKPWK 18
 RESULT 25
 AAB13428
 ID AAB13428 standard; peptide; 19 AA.
 XX
 AC AAB13428;
 XX
 DT 23-NOV-2000 (first entry)
 XX
 DE Synthetic alpha smooth muscle actin inhibitor # 8.
 XX
 KW Alpha smooth muscle actin; alpha-SM; wound contraction;
 KW hypertrophic scar; fibromatosis; Dupuytren disease; fibrotic condition;
 KW lung fibrosis.
 XX
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "Biotinylated Gly"
 FT Modified-site 3
 FT /note= "Tetrapeptide of Acetyl-EEED linked on the epsilon
 FT -amino group of lysyl residue"
 FT

```

XX WO200038733-A1.
PN
XX
XX PD
XX 06-JUL-2000.
XX PF 15-DEC-1999; 99WO-EP009964.
XX PR 24-DEC-1998; 98EP-00204396.
XX (UNIO ) UCB-BIOPRODUCTS SA.
XX PA
XX Gabbiani G, Scarso A;
PI
XX WPI; 2000-452308/39.
XX DR
XX
XX A peptidic product for prevention and treatment of a disease related to
PT alpha-SM actin expression comprises a tetrapeptide associated with a
PT chemical entity that is able to introduce the tetrapeptide into the cell.
XX
XX Claim 7; Page 24; 31pp; English.
XX
XX The present invention relates to novel peptides comprising of a specific
CC tetrapeptide associated with an oligopeptide which allows the
CC introduction of the tetrapeptide into the target cell. The present
CC sequence is one such peptide. The side chain linked to the lysine residue
CC at position 3 corresponds to the specific tetrapeptide, while residues 3
CC to 19 correspond to the oligopeptide. The specific tetrapeptide of the
CC present sequence interferes with alpha smooth muscle (alpha-SM) actin
CC organisation in stress fibers. The present sequence may be used in the
CC prevention and/or treatment of a disease related to alpha-SM actin
CC expression, e.g. wound contraction, hypertrophic scars, fibromatosis and
CC Dupuytren disease and lung fibrosis. The present sequence may also be used to treat
CC invention is described in AAB14263
XX
XX Sequence 19 AA;
SQ

```

```

Query Match      58.1%; Score 75; DB 3; Length 19;
Best Local Similarity 81.2%; Pred. No. 0.00077;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 6 RQPKIWFNRRKPKWK 21
   |||||
Db 4 RQIKIWFNRRMKKK 19

```

Search completed: May 24, 2004, 17:22:22
Job time : 56.1892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 17:18:06 ; Search time 12.2027 Seconds
(without alignments)
165.539 Million cell updates/sec

Title: US-09-977-349-4

Perfect score: 129

Sequence: 1 CSSIRQKIQWPPRRKPWK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR_78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	55.8	234	1 S02014	homeotic protein H
2	69	53.5	48	2 I51439	homeobox protein -
3	69	53.5	76	2 C43559	homeotic protein R
4	69	53.5	86	2 S08303	homeotic protein H
5	69	53.5	245	1 S10032	homeotic protein H
6	69	53.5	264	1 WJHUS2	homeotic protein H
7	69	53.5	264	1 S35219	homeotic protein H
8	68	52.7	60	2 PC2399	antennapedia-like
9	68	52.7	60	2 PC2400	antennapedia-like
10	68	52.7	60	2 I51437	homeotic protein H
11	68	52.7	62	2 B32391	homeotic protein H
12	68	52.7	66	2 S15534	homeotic protein H
13	68	52.7	67	2 B27471	homeotic protein R
14	68	52.7	69	2 S13785	homeotic protein m
15	68	52.7	70	2 B37042	homeotic protein H
16	68	52.7	81	2 A25198	homeotic protein H
17	68	52.7	85	2 T65138	homeotic protein H
18	68	52.7	85	2 A25472	homeotic protein H
19	68	52.7	95	2 B32830	homeotic protein H
20	68	52.7	96	2 A55278	homeotic protein H
21	68	52.7	99	2 A41605	homeotic protein H
22	68	52.7	104	2 D43559	homeotic protein R
23	68	52.7	114	2 A43559	homeotic protein R
24	68	52.7	155	2 S25846	homeotic protein H
25	68	52.7	188	2 I50145	homeotic protein H
26	68	52.7	242	1 WJMSX3	homeotic protein H
27	68	52.7	243	1 WJMS24	homeotic protein H
28	68	52.7	285	1 A43556	homeotic protein H
29	68	52.7	289	2 A43562	homeotic protein H

30	68	52.7	309	1 S09257	homeotic protein H
31	68	52.7	320	1 A39724	homeotic protein H
32	67	51.9	33	2 S57235	antennapedia prote
33	67	51.9	42	2 I65241	homeotic protein H
34	67	51.9	45	2 PC1216	homeotic protein D
35	67	51.9	66	2 S15536	homeotic protein H
36	67	51.9	66	2 S15538	homeotic protein H
37	67	51.9	71	2 JC1161	homeotic protein 3
38	67	51.9	71	2 A60084	homeotic protein H
39	67	51.9	74	2 D34510	homeotic protein H
40	67	51.9	75	2 I51341	homeo box protein
41	67	51.9	75	2 S58852	homeotic protein S
42	67	51.9	78	2 I51342	homeo box protein
43	67	51.9	81	2 S47605	homeotic protein H
44	67	51.9	81	2 B29585	homeotic protein H
45	67	51.9	82	2 S08302	homeotic protein H
46	67	51.9	83	2 S47603	homeotic protein H
47	67	51.9	83	2 S50066	homeotic protein H
48	67	51.9	86	2 A34510	homeotic protein H
49	67	51.9	86	2 J00489	homeotic protein Z
50	67	51.9	87	2 S00589	homeotic protein H
51	67	51.9	88	2 A03317	homeotic protein M
52	67	51.9	96	2 S08639	homeotic protein H
53	67	51.9	96	2 A05266	homeotic protein H
54	67	51.9	97	2 C27176	homeotic protein H
55	67	51.9	97	2 A24779	homeotic protein m
56	67	51.9	103	2 A32167	homeotic protein H
57	67	51.9	105	2 S47602	homeotic protein H
58	67	51.9	105	2 A24771	homeotic protein K
59	67	51.9	106	2 S36448	homeotic protein s
60	67	51.9	107	2 B61045	homeotic protein T
61	67	51.9	113	2 T10775	homeobox protein -
62	67	51.9	118	2 A24777	homeobox protein H
63	67	51.9	118	2 JT0273	homeotic protein H
64	67	51.9	118	2 B24777	homeotic protein M
65	67	51.9	119	2 A03314	homeotic protein m
66	67	51.9	138	2 S20087	homeotic protein b
67	67	51.9	148	2 PC4071	homeobox A5 protei
68	67	51.9	153	1 WJHUS3C	homeotic protein H
69	67	51.9	153	1 WJMSX6	homeotic protein H
70	67	51.9	158	2 A27348	homeotic protein H
71	67	51.9	209	2 A43553	homeotic protein H
72	67	51.9	217	1 WJHUS2C	homeotic protein H
73	67	51.9	217	1 WJMSX2	homeotic protein H
74	67	51.9	220	2 A37371	homeotic protein H
75	67	51.9	220	2 S01063	Hox 2 protein type
76	67	51.9	222	2 S20029	homeotic protein H
77	67	51.9	224	2 S26400	homeotic protein H
78	67	51.9	227	2 S50067	homeotic protein H
79	67	51.9	227	2 A31324	homeotic protein H
80	67	51.9	228	2 S32563	homeotic protein H
81	67	51.9	229	2 A28329	homeotic protein H
82	67	51.9	230	2 S00592	homeotic protein H
83	67	51.9	232	1 A25108	homeotic protein H
84	67	51.9	232	2 S48125	homeotic protein H
85	67	51.9	233	2 I65197	hox1.3 protein - r
86	67	51.9	234	2 S00992	homeotic protein H
87	67	51.9	235	2 A56568	homeotic protein H
88	67	51.9	235	2 S72429	homeotic protein H
89	67	51.9	236	1 S09256	homeotic protein H
90	67	51.9	242	1 A39164	homeotic protein H
91	67	51.9	246	2 T46446	hypothetical prote
92	67	51.9	250	1 A36170	homeotic protein H
93	67	51.9	250	1 A31757	homeotic protein H
94	67	51.9	251	1 B60492	homeotic protein H
95	67	51.9	255	1 WJHUS4B	homeotic protein H
96	67	51.9	259	1 A43551	homeotic protein H
97	67	51.9	269	2 A45578	homeotic protein H
98	67	51.9	270	1 WJHUS1C	homeotic protein H
99	67	51.9	270	1 WJMS13	homeotic protein H
100	67	51.9	275	1 WJZFX2	homeotic protein H

ALIGNMENTS

RESULT 1

S02014
 homeotic protein Hox 1 - eastern newt
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: S02014; S07819
 R:Savard, P.; Gates, P.B.; Brockes, J.P.
 EMBO J. 7, 4275-4282, 1988

A:Title: Position dependent expression of a homeobox gene transcript in relation to amphibian limb development
 A:Reference number: S02014; MUID:89210814; PMID:2907476
 A:Accession: S02014
 A:Molecule type: DNA

A:Residues: 1-234 <SAV>
 A:Cross-references: EMBL:X13957; NID:g64117; PIDN:CAA32139.1; PID:g64118
 R:Tabin, C.J.

Development 105, 813-820, 1989
 A:Title: Isolation of potential vertebrate limb-identity genes.
 A:Reference number: S07819; MUID:90091821; PMID:2574663

A:Accession: S07819

A:Molecule type: mRNA

A:Residues: 131-234 <TAB>

A:Cross-references: EMBL:X16848; NID:g64115; PIDN:CAA34745.1; PID:g64116

C:Genetics:

A:Gene: Hox 1

C:Superfamily: homeotic protein Hox C6; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:1-2/Domain: homeobox homology <HOX>

Query Match 55.8%; Score 72; DB 1; Length 234;
 Best Local Similarity 63.6%; Pred. No. 0.0069; 1;
 Matches 14; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 2 SSCI--RQPKWFPNRRKPWK 21

Db 177 ASCLTERQIKWQNRMRMKWK 198

RESULT 2

I51439
 homeobox protein - African clawed frog (fragment)
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 15-Oct-1999
 C:Accession: I51439
 R:Leroy, P.; DeRobertis, E.M.
 Dev. Dyn. 194, 21-32, 1992

A:Title: Effects of lithium chloride and retinoic acid on the expression of genes from the developing hindbrain
 A:Reference number: I51439; MUID:93043517; PMID:1384809
 A:Accession: I51439

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-48 <LER>

A:Cross-references: GB:M91587; NID:g214257; PIDN:AAA49750.1; PID:g214258

C:Genetics:

A:Gene: Hox2.2

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:1-2/Domain: homeobox homology (fragment) <HOX>

Query Match 53.5%; Score 69; DB 2; Length 48;
 Best Local Similarity 63.6%; Pred. No. 0.0049; 1;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKWFPNRRKPWK 21

Db 6 SLCLTERQIKWQNRMRMKWK 27

RESULT 3

C43559
 homeotic protein R3 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 24-Sep-1999
 C:Accession: C43559
 R:Falzon, M.; Chung, S.Y.
 Development 103, 601-610, 1988

A:Title: The expression of rat homeobox-containing genes is developmentally regulated and

A:Reference number: A43559; MUID:89231502; PMID:2907739

A:Accession: C43559

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-76 <FAL>

A:Cross-references: GB:M37567; NID:g204634; PIDN:AAA41343.1; PID:g204635

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:20-76/Domain: homeobox homology (fragment) <HOX>

Query Match 53.5%; Score 69; DB 2; Length 76;
 Best Local Similarity 63.6%; Pred. No. 0.0072; 1;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKWFPNRRKPWK 21

Db 47 SLCLSERQIKWQNRMRMKWK 68

RESULT 4

S08303
 homeotic protein Hox 2.2 - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C>Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 17-Oct-1997
 C:Accession: S08303
 R:Wedden, S.E.; Pang, K.; Eichele, G.
 Development 105, 639-650, 1989
 A:Title: Expression pattern of homeobox-containing genes during chick embryogenesis.
 A:Reference number: S08302; MUID:90126373; PMID:2575515
 A:Accession: S08303

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-86 <WED>

A:Cross-references: EMBL:X16847

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:10-66/Domain: homeobox homology <HOX>

Query Match 53.5%; Score 69; DB 2; Length 86;
 Best Local Similarity 63.6%; Pred. No. 0.0079; 1;
 Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 2 SSCI--RQPKWFPNRRKPWK 21

Db 45 SLCLTERQIKWQNRMRMKWK 66

RESULT 5

S10092
 homeotic protein Hox B4 - chicken
 N:Alternate names: homeotic protein Chox-2.6; homeotic protein Chox-Z
 C:Species: Gallus gallus (chicken)
 C>Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
 C:Accession: S10092
 R:Sasaki, H.; Kuroiwa, A.
 Nucleic Acids Res. 18, 184, 1990

A:Title: The nucleotide sequence of the cDNA encoding a chicken deformed family homeobox

A:Reference number: S10092; MUID:90174917; PMID:1968620

A:Accession: S10092

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-245 <SAS>

A:Cross-references: EMBL:X17612; NID:g63504; PIDN:CAA35614.1; PID:g63505

C:Function:

A:Description: control of embryonic development by tissue- and stage-specific regulation

C:Superfamily: homeotic protein Hox D4; homeobox homology

C:Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

F:151-207/Domain: homeobox homology <HOX>

```

Query Match          53.5%; Score 69; DB 1; Length 245;
Best Local Similarity 63.6%; Pred. No. 0.019;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY  2 SSCI--RQPKIWFPPNRRKPWK 21
    |: ||||| |||||
Db   186 SLCLSERQIKIWFQNRKWK 207
    |: ||||| |||||

RESULT 6
WUHUJE
homeotic protein Hox C4 - human
N/Alternate names: homeotic protein cpl9; homeotic protein Hox 3E
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C/Accession: S01030; S15545
R/Simeone, A.; Pannese, M.; Acampora, D.; d'Episito, M.; Boncinelli, E.
Nucleic Acids Res. 16, 5379-5390, 1988
A/Title: At least three human homeoboxes on chromosome 12 belong to the same transcript
A/Reference number: S01030; MUID:88262550; PMID:2898768
A/Accession: S01030
A/Molecule type: mRNA
A/Residues: 1-264 <SIM>
A/Cross-references: EMBL:X07495; NID:G32385; PIDN:CAA30376.1; PID:G32386
R/Boncinelli, E.; Acampora, D.; Pannese, M.; d'Episito, M.; Somma, R.; Gaudino, G.; Stc
Genome 31, 745-756, 1989
A/Title: Organization of human class I homeobox genes.
A/Reference number: S15036; MUID:90215256; PMID:2576652
A/Accession: S15545
A/Molecule type: DNA
A/Residues: 156-221 <BON>
C/Genetics:
A/Gene: GDB:HOXC4
A/Cross-references: GDB:120672; OMIM:142974
A/Map position: 12q13.3-12q13.3
C/Function:
A/Description: control of embryonic development by tissue- and stage-specific regulation
C/Superfamily: homeotic protein Hox D4; homeobox homology
C/Keywords: alternative splicing; DNA binding; embryo; homeobox; nucleus; transcription
F:157-213/Domain: homeobox homology <HOX>

Query Match          53.5%; Score 69; DB 1; Length 264;
Best Local Similarity 63.6%; Pred. No. 0.02;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY  2 SSCI--RQPKIWFPPNRRKPWK 21
    |: ||||| |||||
Db   192 SLCLSERQIKIWFQNRKWK 213
    |: ||||| |||||

RESULT 7
S35219
homeotic protein Hox C4 - mouse
N/Alternate names: homeotic protein Hox 3.5; homeotic protein MAB87
C/Species: Mus musculus (house mouse)
C/Date: 10-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C/Accession: S35219; A49153; C41606; I49752
R/Goto, J.; Miyabayashi, T.; Wakamatsu, Y.; Takahashi, N.; Muramatsu, M.
Mol. Gen. Genet. 239, 41-48, 1993
A/Title: Organization and expression of mouse Hox3 cluster genes.
A/Reference number: S35219; MUID:93288004; PMID:8099712
A/Accession: S35219
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-264 <GOT>
A/Cross-references: GB:S62287; NID:G385749; PIDN:AAB27153.1; PID:G385750; GB:D11328; NID
A/Note: entry MUSHOX35A in GenBank release 103 duplicates GenBank entry S62287 except fo
R/Gada, A.M.; Gaunt, S.J.; Azzawi, M.; Shinfeld, S.M.; Pearce, J.; Sharpe, P.T.
Development 116, 497-506, 1992
A/Title: Sequence and embryonic expression of the murine Hox-3.5 gene.

```

```

A/Reference number: A49153; MUID:93161956; PMID:1363091
A/Accession: A49153
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-79; 'G', '81-95', 'S', '97-264 <GEA>
A/Cross-references: GB:X69019; NID:G396183; PIDN:CAA48784.1; PID:G396184
A/Note: sequence extracted from NCHI backbone (NCBI:124829, NCBI:124830)
R/Murtha, M.T.; Leckman, J.F.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 88, 10711-10715, 1991
A/Title: Detection of homeobox genes in development and evolution.
A/Reference number: A41606; MUID:92073357; PMID:1720547
A/Accession: C41606
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 177-201 <MUR>
A/Cross-references: GB:M81660; NID:G193975; PIDN:AAA63313.1; PID:G193976
C/Genetics:
A/Gene: Hoxc-4; Hox 3.5
A/Map position: 15
A/Introns: 147/1
C/Function:
A/Description: control of embryonic development by tissue- and stage-specific regulation
C/Superfamily: homeotic protein Hox D4; homeobox homology
C/Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation
F:157-213/Domain: homeobox homology <HOX>

Query Match          53.5%; Score 69; DB 1; Length 264;
Best Local Similarity 63.6%; Pred. No. 0.02;
Matches 14; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY  2 SSCI--RQPKIWFPPNRRKPWK 21
    |: ||||| |||||
Db   192 SLCLSERQIKIWFQNRKWK 213
    |: ||||| |||||

RESULT 8
PC2399
antennapedia-like homeotic protein AHox 2 - sea squirt (Styela clava) (fragment)
C/Species: Styela clava
C/Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 21-Jul-2000
C/Accession: PC2399
R/Ge, T.; Lee, H.; Tomlinson, C.R.
Gene 147, 219-222, 1994
A/Title: Identification of an antennapedia-like homeobox gene in the ascidians Styela cl
A/Reference number: PC2399; MUID:95011617; PMID:7926803
A/Accession: PC2399
A/Molecule type: DNA
A/Residues: 1-60 <GET>
A/Cross-references: GB:S73920; NID:G693714; PIDN:AAB33061.2; PID:G7387472
A/Note: The authors translated the codon ATT for residue 47 as Glu
C/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:2-58/Domain: homeobox homology <HOX>

Query Match          52.7%; Score 68; DB 2; Length 60;
Best Local Similarity 75.0%; Pred. No. 0.0082;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  6 RQPKIWFPPNRRKPWK 21
    |: ||||| |||||
Db   43 RQVKIWFQNRKWK 58
    |: ||||| |||||

RESULT 9
PC2400
antennapedia-like homeotic protein AHox 3 - sea squirt (Styela plicata) (fragment)
C/Species: Styela plicata
C/Date: 26-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Oct-1997
C/Accession: PC2400
R/Ge, T.; Lee, H.; Tomlinson, C.R.
Gene 147, 219-222, 1994
A/Title: Identification of an antennapedia-like homeobox gene in the ascidians Styela cl
A/Reference number: PC2399; MUID:95011617; PMID:7926803

```